Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 18 (2016), Issue 12 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoints

Original Papers

Viewpoint

Bridging the Health Data Divide

Leo Anthony Celi[1*], MD, MSc, MPH; Guido Davidzon[1*], MD, SM; Alistair EW Johnson[1*], DPhil; Matthieu Komorowski[1*], MD, MRes; Dominic C Marshall[1*], MBBS; Sunil S Nair[1*], MD, MBA; Colin T Phillips[1*], MD; Tom J Pollard[1*], PhD; Jesse D Raffa[1*], PhD; Justin D Salciccioli[1*], MA, MBBS; Francisco Muge Salgueiro[1*], MD; David J Stone[1*], MD

MIT Critical Data, Cambridge, MA, United States
[*]all authors contributed equally

Corresponding Author:
Matthieu Komorowski, MD, MRes
MIT Critical Data
77 Massachusetts Avenue
Cambridge, MA, 02139
United States
Phone: 1 617 710 3114
Fax: 1 617 258 7859
Email: matthieu.komorowski@gmail.com

Abstract

Fundamental quality, safety, and cost problems have not been resolved by the increasing digitization of health care. This digitization has progressed alongside the presence of a persistent divide between clinicians, the domain experts, and the technical experts, such as data scientists. The disconnect between clinicians and data scientists translates into a waste of research and health care resources, slow uptake of innovations, and poorer outcomes than are desirable and achievable. The divide can be narrowed by creating a culture of collaboration between these two disciplines, exemplified by events such as datathons. However, in order to more fully and meaningfully bridge the divide, the infrastructure of medical education, publication, and funding processes must evolve to support and enhance a learning health care system.

(J Med Internet Res 2016;18(12):e325) doi:10.2196/jmir.6400

KEYWORDS
electronic health records; machine learning; health care policy; medical education; collaboration

Introduction

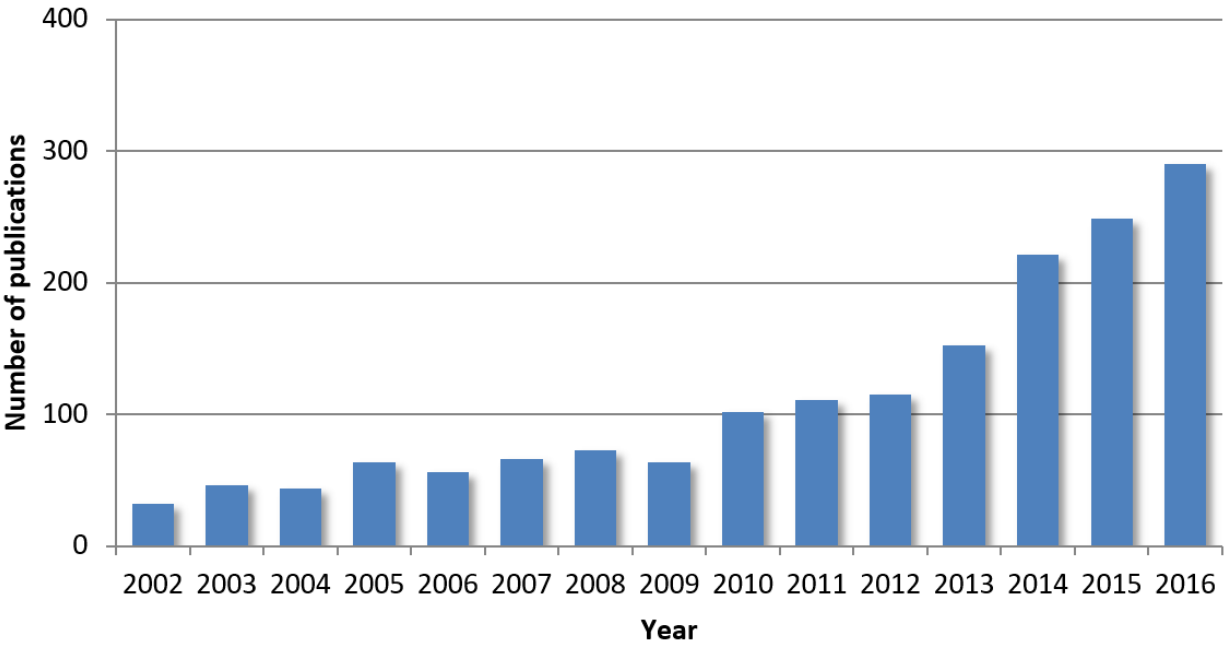
The US Agency for Healthcare Research and Quality (AHRQ) was established in 1989 in response to an Institute of Medicine (now the National Academy of Medicine) report that pointed out “escalating healthcare costs, wide variations in medical practice patterns, and evidence that some health services are of little or no value” [1]. More than 25 years later, there has been surprisingly, perhaps even shockingly, little progress in these three areas. Quality of care, as would be reflected by the universal provision of standardized, evidence-based, and truly indicated care, has not improved to the degree one would have hoped. Similarly, while medical safety and errors have also come increasingly into the awareness of the medical system over that 25-year time period [2], advances in these areas have been slow, hard won, and unsupported by the kinds of smart, data-driven engineering designs that have gone into other domains.

Recent increases in computing power and data storage have resulted in an entirely new field involving the analysis of digitally archived information to acquire new knowledge: data science. While quality of care has largely been defined by clinical trials and expensive prospective studies, the application of data science to the clinical domain has the opportunity to dramatically increase the speed at which knowledge is generated and the breadth of questions that can be answered. The answers given are of particular interest when the research would otherwise be impractical, one such example being the comparison of an augmented treatment with a small effect size (normally requiring a prohibitively large prospective cohort). Data science has the opportunity to inform clinical decision making more directly as well, forecasting the occurrence of relevant clinical phenomena such as physiologic deterioration, diagnosis, medication adherence, or organ rejection. Machine learning, a field that was nascent in 1989 when the AHRQ was established, has now become ubiquitous and informs aspects of our everyday life from search queries to optimal routes. In the decade preceding the publication of the Surviving Sepsis Campaign guidelines [3], there was a significant increase in the number of publications that evaluated the use of machine learning in decision support and prognostication in sepsis (Figure 1).

The interest in applying machine learning to clinical practice is increasing, yet the practical application of these techniques has been less than desirable. Practitioners continue to make determinations in a technically unsupported and unmonitored manner due to a lack of high-quality evidence or tools to support most day-to-day decisions, and as a result the rate of diagnostic errors by individual practitioners is unacceptably high [4].

There is a persistent gap between the clinicians required to understand the clinical relevance of the data and the data scientists who are critical to extracting useable information from the increasing amount of health care data that are being generated. In this paper, we focus on the divide between the data science and health care silos, and posit that the lack of integration is the primary barrier to a data revolution in health care. We first discuss published literature that supports the existence of this divide, and then we present recommendations on how to bridge the gap between practicing clinicians and data scientists.

Figure 1. PubMed search results for ("sepsis"[All Fields] OR "septic"[All Fields]) AND ("machine learning"[All Fields] OR "data analysis"[All Fields] OR "data science"[All Fields] OR "engineering"[All Fields] OR "computing"[All Fields] OR "prediction"[All Fields]) AND ("2002/01/01"[PDAT] : "2016/12/31"[PDAT]).



The Divide

In both the practice and research arenas, there exists a divide between scientists and engineers on the one hand and, on the other, the clinicians who are most familiar with the exigencies and uncertainties that define and constrain the practice of medicine. There are several reasons for the development of this situation. Until recently, there have simply been very limited sources of data available to perform research such as the determination of comparative effectiveness, cost analysis, and the elucidation of treatment effect heterogeneity. Further, there has been a lack of strong motivation (ie, little, to no, to negative financial incentives; a lack of substantiating research; and strong individual and industry inertia and resistance) to reduce wasteful and sometimes harmful practice variation, or avoid (generally well-reimbursed) health services that are of little or no value.

In 2012, Kiri Wagstaff of California Institute of Technology’s Jet Propulsion Laboratory published an insightful article entitled “Machine learning that matters,” which pointed out how disconnected much of machine learning research is from important (and real) problems in society, including those of health care [5]. With the aim of refocusing efforts on topics that matter, Wagstaff proposed several “Impact Challenges” that tie machine learning to real-world outcomes, such as saving a human life or making significant financial savings through improved decision making. In a 2014 special issue of the journal *Machine Learning* [6], Rudin and Wagstaff explored the connection between machine learning research and its broader-world applications in more detail, and explicitly emphasized the importance of interdisciplinary collaboration for development of impactful research. While there is a clear need for machine learning in a variety of practical applications, the authors suggested that lack of enthusiasm in top venues to promote such work creates a “contradictory situation” that holds it back. Such situations serve to reinforce the health data divide.

Bridging the Divide

Recommendation 1: Collaboration

Given the definition of the problem, the most obvious solution, and yet the most challenging one, is encouraging collaboration between data scientists and clinicians. The incorporation of statistics into clinical research in the past 50 years and the rise of the biostatistician can act as a template from which the community can learn. The role of a biostatistician in biomedical research has become collaborative over time, at least partly due to the history and traditions of the discipline. An important event in this history in the United States was the passage of the 1962 Kefauver-Harris amendments to the Federal Food, Drug, and Cosmetic Act, which established in a preliminary form the method by which drugs are evaluated by the Food and Drug Administration today [7]. The act itself mandated proof of efficacy for new drugs, which had not been required before its signing into law. This piece of legislation and the ensuing events in the following two decades brought statisticians into close contact with clinician investigators, who now needed the statistician's expertise to design, analyze, and report their study, and effectively established the tradition of collaboration between biostatisticians and clinician scientists seen to this day.

Can this process be accelerated between data scientists and those immersed in the practice of medicine? At the novel Icahn Institute for Genomics and Multiscale Biology at Mount Sinai Health System in New York, more than 300 people were hired to staff the new institute, with backgrounds across hardware design, big data computing, gene sequencing, and bioinformatics [8]. By linking this talent with disease centers within Mount Sinai, and using the tools of machine learning and predictive modeling (elements of big data), scientists have already published on inflammatory models in common-variant Alzheimer disease [9] and are taking a closer look at one of the most complex and biodiverse cell populations in the human body, the gut microbiome, which may be responsible for far more of the body's homeostasis than previously realized [10].

Bridging the divide may be facilitated by instilling researchers with a greater appreciation of the benefits offered through collaboration with colleagues of complementary disciplines. Two papers published in *Advances in Physiology Education*, a journal of the American Physiological Society, call for changes in medical education to do so, with the aim of closing the knowledge gap between engineers and physicians [11,12].

Recommendation 2: Education

Medicine has clumsily entered its digital age via the back door: vast and costly electronic medical records systems have been implemented largely without careful and planned consideration for their impact on the entire health care system, including education, practice, workflows, and research. Education and practice systems have not taken this new digitized world into full account, and consequences include students who are unprepared for their digital futures, very unhappy physicians stuck behind computer screens selecting seemingly endless items in reams of dropdown lists, and the unconscious loss of many opportunities for improvements in practice and research. It is time, even if a bit tardy and somewhat less than proactive, to acknowledge and address this transition of medicine from paper to computer, from opinion and experience to evidence, and from memory to search engines.

Previous reports have demonstrated a deficit in knowledge of clinicians and even clinician scientists relating to statistical methods and their applications to clinical data, including the data used in practice [13,14]. There appears to be little knowledge of, and seemingly even less interest in, these increasingly critical issues among most physicians. This unacceptable awareness and training gap has prompted updates internationally in medical curricula in order to include additional instruction in and exposure to statistical applications and epidemiology. However, there are inadequate numbers of physician educators who are equipped with the knowledge in informatics and data science required to provide even the most basic and essential insights to junior trainees [15].

The questions then consist of who, when, how, and what do we train? Training should focus on two groups of medical trainees: medical students and residents. If we accept that over the next half century there will likely be an increasing need for hybrid skills of this nature, then there is a strong case for inclusion of data science in the core curriculum in medical school and during residency training [16]. An introduction to the use of digital health records for research may provide a foundation to be able to contribute to knowledge discovery regardless of the career path medical students and residents eventually choose. This should then be followed by optional courses, preferably with practical research (eg, summer courses or internships), to further develop these skills for those particularly interested. This latter group is likely to form the core of future educators in this area.

Interested medical students and residents would benefit from educational opportunities that foster cross-disciplinary working relationships. One such experience might be the participation in the datathons that we have previously used in our own work to encourage collaboration [17]. The resources and insights generated during such events may be stored and used in ongoing collaborations and may be continually updated to provide greater scientific rigor and insight. A final suggestion is represented by online platforms and communities, where physicians and data scientists could interact, discuss clinically relevant questions, and share repositories of code and worked examples.

Perhaps most important, creating a medical culture that is aware of and respectful of the importance and potential power of data for supporting and improving both practice and research may be the most important and ultimately effective element. It is desirable that each participant in the clinical process realizes and understands their role in the overall system of providing reliable and robust data that they and others will subsequently use in improving care.

Recommendation 3: Rethinking Academic Incentives

The education of medical trainees in data analysis methods and data scientist trainees in the particular domain issues of clinical practice and data would be a primer for future collaboration between the two groups later in their careers. Such cooperation could be fostered by policies on the part of academic journals that encourage joint submissions from clinicians and data scientists. The perceived "publish or perish" culture of academic medicine has not much changed for the past 30 years [18] and has led to all manner of trivialities: publishing on obscure or irrelevant results with minimal clinical or research importance; sectioning results into multiple manuscripts across numerous journals; and competing for data sets (or, at least, a lack of transparency and sharing of the same). While not a panacea for these ills, equal authorship may ease some of the barriers to cooperation between researchers and perhaps result in higher-quality publications. Big data lends itself better to collaboration than to separation—multiple studies can be run on the same dataset for different purposes. For instance, an emergency department data set from Hong Kong was used both to identify populations who are sensitive to extreme weather, *and* to develop a long-term forecast for ambulance demand through 2036 [19].

We predict that the continued success of such collaborative models will require the recognition of equal authorship as on par with first/last authorship as currently used by journals, universities, and funding agencies. We argue that the impact of the paper is more important than the individual contributions, provided the authors will testify they each had a meaningful role to play in the development of the final publication or proposal. Equal authorship may not end, but could temper, current publishing strategies where multiple, smaller, less impactful papers are released so as to procure first/last authorship for all members of the group—currently important for their curricula vitae and academic standing. A next step would be to give credit for the use of data by counterparties—how better to validate one's cohort than to have other researchers use the same data for their own related (or unrelated) research questions? Any conclusions reached by the latter should in part be attributed to the former for their role in generating and collating a high-fidelity data set, and that role also recognized in academic circles. These measures would encourage investment in big data infrastructure while also improving the quality of "big data" conclusions.

Recommendation 4: Funding

Fostering research in a specific field and encouraging collaboration between fields through funding is not new. Almost 20 years ago, the US National Institutes of Health (NIH) convened a Director's Panel on Clinical Research to address the decline in physician investigators applying for clinical research grants. The panel proposed a series of recommendations to increase the funding opportunities for clinical researchers [20]. We suggest implementing many of the same methods today to encourage closer collaboration between clinicians and data scientists through funding and incentives.

NIH-funded research programs are classified using activity codes, with, for example, codes from the R series corresponding to research grants (eg, R01, R13) and from the K series corresponding to careers development awards. We suggest creating a K award category for clinicians engaged in data science similar to a K23 grant for clinicians that the Director's Panel on Clinical Research proposed. Providing a unique K grant for data scientists would provide increased funding opportunities, since K awards typically fund 40% of applicants, unlike the R0 awards, which are more competitive and funded at a much lower rate, usually 10%. This support is critical to foster data science investigators during the vulnerable early period in their career.

By restructuring study sections and adding study sections across all the institutes exclusive to secondary analysis, grants would increase the funding available for data science proposals that may not otherwise obtain funding in the current structure. Having a study section specific to secondary analysis to score data science proposals would ensure that at least 10% of the data science proposals are funded (and avoid competing against primary analysis proposals, which might receive preferential scoring). In addition, mandating a health care provider as part of the proposal team would encourage the clinical impetus behind the proposal.

Another example of the US federal government increasing access to funding is the US Department of Veterans Affairs' Big Data Scientist Training Enhancement Program, which has been adopted by 6 pilot centers [21]. The program supports data scientists working directly with clinician scientists on-site at a hospital.

Conclusion

Better use of clinical data has the potential to address a number of important, problematic, and unresolved issues in the health care system. These include high, and perhaps excessive, costs; unnecessary and undesirable practice variation; the improvement of digital workflows; the universal implementation of a reasonable, reliable, and usable version of evidence-based medicine; the introduction of personalized and precision medicine; quality; safety; effective communication; efficient care coordination; and the introduction of data-driven and -supported clinical decision making. However, the introduction of this kind of revolution into health care inevitably involves crossing disciplinary boundaries in a way that requires cooperation and collaboration among a frankly diverse group of experts in order to optimize the combined output of these contributors. The formation of such teams requires that each team member be more educated in the issues involved outside of their own comfort areas. As a primarily medically oriented group, we focus on the impact on medical training, but the principles relate to those in nonmedical areas who need to become sufficiently educated in clinical matters to contribute optimally to the grand scheme. For example, how can current advanced analytic techniques such as machine learning be best applied to both clinical research and practice problems? Clearly, specific kinds of clinical-technical collaborations will be required to guide these kinds of processes and projects to fruition.

In this paper, we have attempted to portray not just the problems, but also potential solutions, or at least beginning approaches to solutions, for the situation in which we find ourselves. This situation involves a costly, complex, and massive health care system that can well bear improvement, and a growing mound of underutilized data that is accumulating as a result of the accelerating digitization of medical care.

Clinicians should not feel like interchangeable cogs entering reams of data blindly into a vast black hole of no return; data scientists should not be discovering new knowledge and developing predictive algorithms isolated from the domain experts. Rather, all should see themselves as diversely necessary components of a truly functional clinical data system that works toward providing excellent care to individuals and populations while working to improve all facets of that care.

Take-Home Messages

1. Fundamental quality, safety, and cost problems have not been resolved by the increasing digitization of health care.

2. This digitization has progressed alongside the presence of a persistent divide between clinicians, the domain experts, and the technical experts, such as data scientists.

3. The divide can be narrowed by creating a culture of collaboration between these two disciplines, exemplified by events such as datathons.

4. However, in order to more fully and meaningfully bridge the divide, the infrastructure of medical education, publication, and funding processes must evolve to support and enhance a learning health care system.

Acknowledgments

We thank Dr Roger Mark, Dr Kenneth Mukamal, and Dr Peter Weller for their invaluable input.

Leo Anthony Celi, Alistair EW Johnson, Tom J Pollard, and Jesse Raffa are funded by the National Institutes of Health through National Institute of Biomedical Imaging and Bioengineering grant R01 EB017205-01A1 and by Philips Healthcare. Matthieu Komorowski is funded by an Imperial College London President’s PhD scholarship.

Conflicts of Interest

None declared.

References

1. Grossman JH, Field MJ, Lohr KN. Clinical Practice Guidelines: Directions for a New Program. Washington, DC: National Academy Press; 1990.
2. Institute of Medicine, Committee on Quality of Health Care in America. In: Kohn LT, Donaldson MS, Corrigan JM, editors. To Err Is Human: Building a Safer Health System. Washington, DC: National Academies Press; 2000.
3. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, Surviving Sepsis Campaign Guidelines Com. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock, 2012. Intensive Care Med 2013 Feb;39(2):165-228. [doi: 10.1007/s00134-012-2769-8] [Medline: 23361625]
4. Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf 2013 Oct;22 Suppl 2:ii21-ii27 [FREE Full text] [doi: 10.1136/bmjqs-2012-001615] [Medline: 23771902]
5. Wagstaff K. Machine learning that matters. Ithaca, NY: Cornell University Library; 2012. URL: https://arxiv.org/ftp/arxiv/papers/1206/1206.4656.pdf [accessed 2016-12-08] [WebCite Cache ID 6mZzcqjHR]
6. Rudin C, Wagstaff KL. Machine learning for science and society. Machine Learning 2013 Nov 28;95(1):1-9. [doi: 10.1007/s10994-013-5425-9]
7. Greene JA, Podolsky SH. Reform, regulation, and pharmaceuticals--the Kefauver-Harris Amendments at 50. N Engl J Med 2012 Oct 18;367(16):1481-1483 [FREE Full text] [doi: 10.1056/NEJMp1210007] [Medline: 23075174]
8. Schadt E, Chilukuri S. The role of big data in medicine. New York, NY: McKinsey & Company; 2015. URL: http://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/the-role-of-big-data-in-medicine [accessed 2016-12-08] [WebCite Cache ID 6jGWtN6jt]
9. Zhang B, Gaiteri C, Bodea L, Wang Z, McElwee J, Podtelezhnikov AA, et al. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. Cell 2013 Apr 25;153(3):707-720 [FREE Full text] [doi: 10.1016/j.cell.2013.03.030] [Medline: 23622250]
10. Icahn Institute and Department of Genetics and Genomic Sciences. Research. New York, NY: Icahn School of Medicine at Mount Sinai; 2016. URL: http://icahn.mssm.edu/research/genomics/research [accessed 2016-12-08] [WebCite Cache ID 6jGX30J9V]
11. Humphrey JD, Coté GL, Walton JR, Meininger GA, Laine GA. A new paradigm for graduate research and training in the biomedical sciences and engineering. Adv Physiol Educ 2005 Jun;29(2):98-102 [FREE Full text] [doi: 10.1152/advan.00053.2004] [Medline: 15905153]
12. Matsuki N, Takeda M, Yamano M, Imai Y, Ishikawa T, Yamaguchi T. Effects of unique biomedical education programs for engineers: REDEEM and ESTEEM projects. Adv Physiol Educ 2009 Jun;33(2):91-97 [FREE Full text] [doi: 10.1152/advan.90120.2008] [Medline: 19509393]
13. Horton NJ, Switzer SS. Statistical methods in the journal. N Engl J Med 2005 Nov 03;353(18):1977-1979. [doi: 10.1056/NEJM200511033531823] [Medline: 16267336]
14. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. JAMA 2007 Sep 05;298(9):1010-1022. [doi: 10.1001/jama.298.9.1010] [Medline: 17785646]

15. Lucey CR. Medical education: part of the problem and part of the solution. JAMA Intern Med 2013 Sep 23;173(17):1639-1643. [doi: 10.1001/jamainternmed.2013.9074] [Medline: 23857567]
16. Moskowitz A, McSparron J, Stone DJ, Celi LA. Preparing a new generation of clinicians for the era of big data. Harv Med Stud Rev 2015 Jan;2(1):24-27 [FREE Full text] [Medline: 25688383]
17. Aboab J, Celi LA, Charlton P, Feng M, Ghassemi M, Marshall DC, et al. A “datathon” model to support cross-disciplinary collaboration. Sci Transl Med 2016;8:333ps8. [doi: 10.1126/scitranslmed.aad9072] [Medline: 27053770]
18. Angell M. Publish or perish: a proposal. Ann Intern Med 1986 Feb;104(2):261-262. [Medline: 3946958]
19. Wong HT, Yin Q, Guo YQ, Murray K, Zhou DH, Slade D. Big data as a new approach in emergency medicine research. J Acute Dis 2015 Aug;4(3):178-179. [doi: 10.1016/j.joad.2015.04.003]
20. Nathan DG. Clinical research: perceptions, reality, and proposed solutions. National Institutes of Health Director's Panel on Clinical Research. JAMA 1998 Oct 28;280(16):1427-1431. [Medline: 9801003]
21. Barbero R, Levy N. Increasing the use of data science in U.S. Department of Veterans Affairs Medical Centers. Washington, DC: The White House; 2015. URL: https://www.whitehouse.gov/blog/2015/10/08/increasing-use-data-science-us-department-veterans-affairs-medical-centers [accessed 2016-12-08] [WebCite Cache ID 6jGX4T93R]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
NIH: National Institutes of Health

Edited by G Eysenbach; submitted 28.07.16; peer-reviewed by K Wagholikar, K Jiggins Colorafi; comments to author 28.09.16; revised version received 28.09.16; accepted 21.11.16; published 20.12.16.

Please cite as:
Celi LA, Davidzon G, Johnson AEW, Komorowski M, Marshall DC, Nair SS, Phillips CT, Pollard TJ, Raffa JD, Salciccioli JD, Salgueiro FM, Stone DJ
Bridging the Health Data Divide
J Med Internet Res 2016;18(12):e325
URL: http://www.jmir.org/2016/12/e325/
doi:10.2196/jmir.6400
PMID:27998877

Original Paper

The Effectiveness of Web-Based Asthma Self-Management System, My Asthma Portal (MAP): A Pilot Randomized Controlled Trial

Sara Ahmed[1,2,3,4], BPhysio, MSc (Rehab), PhD; Pierre Ernst[5], FRCPC, MD; Susan J Bartlett[1,3], PhD; Marie-France Valois[3], MSc; Tasneem Zaihra[1,2,4,6], PhD; Guy Paré[7], PhD; Roland Grad[8], MSc, FCFP, MDCM; Owis Eilayyan[1,2,4], MSc; Robert Perreault[9], FRCPC, MD; Robyn Tamblyn[3,10], PhD

[1]Division of Clinical Epidemiology, McGill University Health Center, McGill University, Montreal, QC, Canada
[2]School of Physical & Occupational Therapy, Faculty of Medicine, McGill University, Montreal, QC, Canada
[3]Department of Medicine, Faculty of Medicine, McGill University, Montreal, QC, Canada
[4]Centre de recherche interdisciplinaire en réadaptation (CRIR), Constance Lethbridge Rehabilitation Center, Montreal, QC, Canada
[5]Centre for Clinical Epidemiology, Lady Davis Research Institute, Jewish General Hospital, McGill University, Montreal, QC, Canada
[6]The College at Brockport, Department of Mathematics, State University of New York, Brockport,, NY, United States
[7]Department of Information Technology, HEC Montreal, Montreal, QC, Canada
[8]Department of Family Medicine, Faculty of Medicine, McGill University, Montreal, QC, Canada
[9]Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, QC, Canada
[10]Biostatistics and Occupational Health and the Clinical and Health Informatics Research Group, Department of Epidemiology, McGill University, Montreal, QC, Canada

Corresponding Author:
Sara Ahmed, BPhysio, MSc (Rehab), PhD
McGill University Health Center
Division of Clinical Epidemiology
McGill University
3654 Prom Sir-William-Osler
Montreal, QC, H3G 1Y5
Canada
Phone: 1 514 398 4400 ext 00531
Fax: 1 514 398 6360
Email: sara.ahmed@mcgill.ca

Abstract

Background: Whether Web-based technologies can improve disease self-management is uncertain. My Asthma Portal (MAP) is a Web-based self-management support system that couples evidence-based behavioral change components (self-monitoring of symptoms, physical activity, and medication adherence) with real-time monitoring, feedback, and support from a nurse case manager.

Objective: The aim of this study was to compare the impact of access to a Web-based asthma self-management patient portal linked to a case-management system (MAP) over 6 months compared with usual care on asthma control and quality of life.

Methods: A multicenter, parallel, 2-arm, pilot, randomized controlled trial was conducted with 100 adults with confirmed diagnosis of asthma from 2 specialty clinics. Asthma control was measured using an algorithm based on overuse of fast-acting bronchodilators and emergency department visits, and asthma-related quality of life was assessed using the Mini-Asthma Quality of Life Questionnaire (MAQLQ). Secondary mediating outcomes included asthma symptoms, depressive symptoms, self-efficacy, and beliefs about medication. Process evaluations were also included.

Results: A total of 49 individuals were randomized to MAP and 51 to usual care. Compared with usual care, participants in the intervention group reported significantly higher asthma quality of life (mean change 0.61, 95% CI 0.03 to 1.19), and the change in asthma quality of life for the intervention group between baseline and 3 months (mean change 0.66, 95% CI 0.35 to 0.98) was not seen in the control group. No significant differences in asthma quality of life were found between the intervention and control groups at 6 (mean change 0.46, 95% CI –0.12 to 1.05) and 9 months (mean change 0.39, 95% CI –0.2 to 0.98). For poor control status, there was no significant effect of group, time, or group by time. For all self-reported measures, the intervention group had a significantly higher proportion of individuals, demonstrating a minimal clinically meaningful improvement compared with the usual care group.

Conclusions: This study supported the use of MAP to enhance asthma quality of life but not asthma control as measured by an administrative database. Implementation of MAP beyond 6 months with tailored protocols for monitoring symptoms and health behaviors as individuals' knowledge and self-management skills improve may result in long-term gains in asthma control.

ClinicalTrial: International Standard Randomized Controlled Trial Number (ISRCTN): 34326236; http://www.isrctn.com/ISRCTN34326236 (Archived by Webcite at http://www.webcitation.org/6mGxoI1R7).

(J Med Internet Res 2016;18(12):e313) doi:10.2196/jmir.5866

KEYWORDS

Internet; nursing; case management; self-care; asthma; quality of life

Introduction

The episodic nature of asthma symptoms and exacerbations makes effective self-management of the condition an imperative. Ongoing self-monitoring of asthma control coupled with an individualized written asthma action plan is required to allow patients to quickly identify and address poor disease control and mild exacerbations by increasing medication or contacting a physician [1]. However, less than 40% of individuals with asthma regularly monitor their symptoms [2] and even fewer initiate their prescribed action plan at the first signs of an exacerbation [3,4]. Further, ongoing self-monitoring and self-management is in part related to the extent to which patients' treating physician and care team provide review and feedback [5-7]. In addition, many patients have difficulty recognizing signs of early deterioration and often fail to utilize their asthma action plans and access their physicians in a timely manner to control worsening asthma status [8,9].

Providing ongoing communication and self-management support is challenging given limited clinical resources and time [10,11]. Health information technology may offer unique opportunities to facilitate integrated self-management support by providing a means for ongoing monitoring and feedback combined with 2-way communication and information exchange between the patient and his or her care team between visits. Providing patients with chronic conditions timely access to health information and personalized alerts, when there is a need for action, may empower individuals to self-manage more effectively by facilitating and reinforcing health behavior change.

A review that investigated the effectiveness of using computers to deliver patient self-management programs for chronic illness found that there was insufficient evidence to determine whether computer-enabled interventions were effective compared with no intervention or usual care [10]. Although the results have been inconsistent, a systematic review of computer-based interventions for asthma supported that these interventions may be effective in improving knowledge, reducing activity limitations, improving markers of self-management, improving quality of life, and optimizing medication use [12]. However, 21% of trials were among children. Most studies failed to report information on the socioeconomic status of participants, and few included individuals over the age of 50 years, limiting generalizability. Furthermore, most trials have only evaluated isolated components required for self-management of asthma, restricting the comprehensiveness of the self-management support received by the patients, and only one study had the support of a case manager to intervene when needed. Finally, whereas patient report is necessary to evaluate the impact of behavioral interventions, previous studies in asthma have not included an evaluation of asthma control independent of participant reporting such as the combined need for emergency care and overuse of rescue medications [13,14]. These outcomes provide necessary information to assess the impact of Web-based interventions on longer-term outcomes, less influenced by daily fluctuations in symptoms, and address the impact of disease on individuals' consumption of health care services and related costs.

More evidence is needed to determine whether Web-based self-management support systems are helpful in clinical practice and can lead to better outcomes. In addition, more information is needed about the proportion of patients who are willing and able to use a Web-based self-management tool and the functionalities that are most frequently accessed. The primary objective of this study was to compare the impact of access to a customized personal Web-based self-management asthma patient portal, My Asthma Portal (MAP), linked to a case-management system over 6 months compared with usual care on asthma control and asthma quality of life. We hypothesized that at 6 months, use of MAP would be associated with improved asthma control and asthma quality of life.

Methods

Ethical Considerations

The McGill University Health Center ethics committee provided approval for this study. Research nurses provided patients with an information brochure, and patients were invited to speak to the research assistant for further information. The research assistant obtained informed patient consent. The detailed protocol is available elsewhere [15].

Setting and Participants

Participants were recruited from pulmonary clinics in 2 tertiary care hospitals located in Montreal, Canada. Participants were eligible if they were aged between 18 and 69 years, had a physician diagnosis of asthma, were prescribed at least one rescue medication, classified as having poor asthma control at the time of recruitment by their treating physician, had access to the Internet, reported smoking <20 pack-years, and were fluent in English or French. We excluded individuals with a diagnosis of chronic obstructive pulmonary disease or other serious medical diagnoses (eg, lung cancer), and those with severe mobility limitations. Potentially eligible participants were identified and approached by their pulmonologist or the nursing staff. The research assistant provided additional information about the study, responded to questions, and obtained written consent.

Study Design

The study was a 6-month, multicenter, parallel, 2-arm, pilot, randomized controlled trial (ISRCTN34326236). Participants were randomly assigned to 1 of the 2 conditions: (1) MAP access or (2) usual care. Randomization occurred at the individual patient level in 2 participating pulmonary clinics. Treatment allocation was done by random permutation within blocks with block sizes of 4 and 6 using a computerized algorithm.

Interventions

Intervention: My Asthma Portal

MAP was developed to support patient self-management and facilitate communication with the care team between visits using an iterative design process based on behavior change and self-efficacy theory [15]. The impact of self-management on health outcomes is thought to occur primarily through changes in health behaviors by helping patients develop the confidence to engage in tasks and acquire core knowledge and skills aimed at helping them better manage their health. Therefore, the features within MAP are based on increasing individuals' knowledge as well as their confidence to carry out the behaviors that are necessary for patients to improve their asthma health. These features are summarized in Table 1 of the protocol publication [15].

Participants were given a MAP username and password, advised that they could access the site from anywhere (eg, home, work, library), and asked to login at least once per week for the 6 months of the study (access to MAP ended after this period). The log-in frequency was decided based on clinician feedback, the fact that symptoms are monitored over 1-week period, and patient feedback on what log-in frequency they believed was feasible as obtained during the development phase. Furthermore, based on a systematic review, the typical Web-based intervention is used once a week [16]. During the first log-in, patients provided basic demographic and health information (eg, smoking status, allergies, and triggers) and selected their own initial health goals for action.

Within MAP, participants could (1) view their personal health information (eg, asthma medications, other health problems); (2) view general asthma information through links to specific educational websites (Learning Center) and receive information tailored to identified knowledge gaps (eg, current medications); and (3) monitor and receive feedback regarding current self-management practices.

Initial feedback was generated automatically from the MAP system using monitoring information entered by the participant and data from the provincial administrative database sent to MAP from the Medical Office for the twenty first century (MOXXI), an electronic health record with a prescription and computerized drug and disease management system. A summary of the logic for color-coding and recommendations offered has been previously summarized [17] and is presented in Multimedia Appendix 1. The color-coding was based on the Canadian asthma guidelines for symptom management [1]. An email alert was sent to patients under the following conditions: (1) when suboptimal asthma control was identified by the system and the participant indicated they did not initiate activities outlined in their action plan; (2) if the action plan was updated by the nurse as prescribed by the patient's physician; and (3) if they had not logged in for at least seven days. The participant was given an opportunity to act on these alerts; if no action had occurred within 48 hours, the alert status was escalated to notify the nurse case manager.

Participants communicated and received support as needed from a nurse case manager using MAP. The nurse case-management system was designed to (1) quickly identify patients that may require immediate care; (2) collate relevant medical and monitoring information for each participant; and (3) document case-management information, including interactions (phone and emails), along with advice and interventions provided by the case manager. Upon receiving this notice, the nurse followed up with the participant within 24 hours through the MAP mail system or by telephone. Additional alerts were sent to the nurse case manager when a new participant was enrolled and needed their prescribed action plan reviewed, the participant indicated lack of understanding of the action plan, or there was a change in medication use based on the self-reported adherence monitoring assessment.

Control Intervention: Usual Care

The control group did not have access to MAP. All participants in the control and intervention groups continued to receive ongoing asthma care from their pulmonologist throughout the trial, and an asthma nurse provided education and follow-up sessions as needed. Topics (similar to those found in the static learning module of MAP) included the importance of avoiding triggers, taking all asthma medications as prescribed, and using the written action plan. The asthma nurse conducted follow-up phone calls between visits, when appropriate (ie, missed appointments, to clarify aspects of the action plan or prescribed asthma medications).

Measures

Process Evaluation

Automated audit trails (computer logs) were used to evaluate MAP usage rates. Use was defined as the frequency and intensity, mainly the number of times and minutes patients spent logged into the system. Patterns of usage included review of the days/week and amount of time that patients used the system. Features used included the number of times and topic of the messages sent to the nurse case manager.

We also evaluated acceptability and attitude of participants toward the Web portal. We used adapted versions of instruments associated with the technology acceptance model (TAM) to assess perceptions of usefulness and ease of use [18-21]. Items were scored from 1 (strongly disagree) to 5 (strongly agree). In other studies, reliability of ratings ranged from 0.92 to 0.98 for usefulness and 0.88 to 0.94 for ease of use [18,20]. Perceived usefulness and ease of use have also been found to be significantly correlated with both self-reported current usage (*r*=.63, *r*=.45, respectively) and self-predicted future usage (*r*=.85, *r*=.59, respectively) [18].

Primary Outcomes

Asthma Control at 6 Months

Asthma control at 6 months was evaluated by examining potential overuse (yes/no) of rescue fast-acting bronchodilators (FABA; ie, beta 2-agonists) based on units dispensed for prescriptions as recorded in the provincial drug database that covers drugs dispensed for individuals who are provincially insured, medical services provided, emergency department (ED) visits, and hospitalizations. Excessive use of FABA was chosen because it is associated with an increased risk of hospitalization and asthma mortality [22]. Based on a previously validated algorithm, asthma was classified as poorly controlled if the sum of the quantity for all FABAs dispensed was >500 doses of salbutamol 100 mcg, 2 inhalations at a time, or the equivalent for other FABAs in the last 6 months of follow-up [1], or the participants visited an ED for a pulmonary-related problem in the last 6 months. To evaluate asthma control at 3 months (halfway), half the amount of FABA was used to calculate overuse and ED visits were evaluated over the 3-month period.

Asthma Quality of Life

The Mini-Asthma Quality of Life Questionnaire (MAQLQ) evaluates symptoms, emotions, exposure to environmental stimuli, and activity limitations [17,23] using a 7-point Likert scale. The MAQLQ was chosen because it is internally consistent (interclass correlation coefficient=0.83), correlates highly with the 32-item version (*r*=.90) and moderately (*r*=.69) with the Asthma Control Questionnaire [23,24], and is sensitive to change (*P*<.001) [23].

Secondary Outcomes and Baseline Variables

Demographic Information

Demographic information such as sex, age, education, and socioeconomic status based on postal code were extracted from the baseline questionnaires. Participants reported on education level at baseline.

The Chronic Disease Self-Efficacy Scale

The Chronic Disease Self-Efficacy Scale [25] has been shown to have adequate psychometric properties in patients with chronic conditions, with internal consistency ranging from .77 to .92 and test-retest reliability .72 to .89 [25,26], and was adapted in this study to assess asthma self-efficacy. Participants rated their level of confidence (1=not confident at all to 10=very confident) with respect to taking medications as prescribed, recognizing and acting on symptoms consistent with deteriorating asthma status, knowing when to initiate their action plan, eating a healthy diet, and participating regularly in physical activity.

The Asthma Control Test

The Asthma Control Test (ACT), a 5-point patient-administered survey for assessing asthma control, evaluates patient perceptions of asthma control [27,28]. Patients rate asthma symptoms over the past 4 weeks as well as their overall level of asthma control. The ACT has been found to be internally consistent (Cronbach alpha=.85) [28] and to have moderate test-retest reliability (intracluster correlation=0.77).

Beliefs About Medicines Questionnaire

Patient’s beliefs about their medicines were evaluated using the Beliefs about Medicines Questionnaire (BMQ) [17,29]. The BMQ assesses beliefs about the “necessity” of prescribed medication for controlling their illness as well as “concerns” about the potential adverse consequences. The necessity-concerns differential score is calculated by subtracting the concerns score from the necessity score (range –20 to 20) [30]. A positive differential score indicates that the patient has stronger beliefs in the necessity of medications compared with concerns and vice versa in the case of a negative score. The 2 BMQ subscales are internally consistent (Cronbach alpha for necessity and concern: 0.87 and 0.78, respectively) [30].

Patient Health Questionnaire

Patient Health Questionnaire (PHQ-9) is a tool based on the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, criteria for identifying depressive symptoms that has been validated and widely used in a number of patient populations, including older adults and patients with asthma [31,32]. A PHQ-9 score of 10 has an 88% sensitivity and 88% specificity for symptoms of major depression. Asthmatics with a PHQ-9 score of greater than or equal to 10 were classified as having symptoms suggestive of moderate or more severe depression.

EuroQol Visual Analogue Scale

Health status was evaluated with the EuroQol visual analogue scale (EQ-VAS). Patients were asked to rate their current health state on a 0 (worst imaginable health state) to 100 (best imaginable health state) scale [33].

Asthma-Related ED Visits or Hospitalizations

Asthma-related ED visits or hospitalizations were extracted from the Quebec provincial insurance database using validated algorithms [34]. Information on the date, type, provider, and location (eg, inpatient, emergency, clinic) for all fee-for-service medical services (~86% of all services) [35] is available for all Quebec residents.

All outcomes were collected at 3 months (halfway) and 6 months (end) with follow-up 9 months later to evaluate persistent effects. Questionnaires were administered in the clinic at baseline and via mail thereafter.

Statistical Analyses

The study was designed to have 80% power (alpha=.05) to detect a 10% difference in proportions of patients whose asthma was classified as out of control. This required a total sample size of 67 participants. To account for attrition and loss to follow-up, an additional 33 subjects were recruited (total 100; 50 per arm). This sample size also allowed us to detect a clinically meaningful difference of 0.5 (SD 0.61) on the MAQLQ scores [17].

Descriptive statistics, including means for continuous variables and proportions for categorical variables, were used to summarize participant characteristics. Analyses were conducted using an intention-to-treat analysis.

A generalized estimating equation model [36] was used to compare the proportions identified as having poor asthma control between groups. The clustering effect was at the level of the patient to account for multiple observations on the same person over time. Mixed-effects linear regression [36] modeling was used to estimate the difference in asthma quality of life between groups. For both analyses, patient was the unit of analysis. An auto-regressive by a degree of 1 (AR(1)) correlation structure was used for the asthma control outcome, and compound symmetry (CS) correlation structure was used for the asthma quality of life outcome to account for multiple observations for the same patient [36]. Both types of analyses were performed testing the effects of time, intervention, and an interaction between time and intervention (as fixed effects in the case of mixed-effects models). Analyses were conducted with SAS version 9.3 (SAS Institute Inc) using PROC GENMOD (binary outcome), and PROC MIXED (continuous outcomes). Models were adjusted for number of asthma medications, sex, age, education level, and self-reported health. Sensitivity analyses were conducted to evaluate the effects of missing data by imputing missing data using the Markov chain Monte Carlo method [37] with relative efficiency of 98% in SAS (PROC MI, PROC MIANALYZE).

Secondary analyses were conducted to evaluate the mediating effects of depressive symptoms (PHQ), self-efficacy, beliefs about asthma medication (BMQ), and asthma control (ACT). In addition, the MAQLQ was added to the model with poorly controlled asthma as the outcome and asthma control status for the model with MAQLQ as the outcome.

Furthermore, the proportion of individuals in each group who achieved a minimal clinically important difference (MCID) between follow-up evaluations on self-report measures was also examined. The estimated MCID of the MAQLQ was 0.5 [2,3], the EQ-VAS was 5, and the PHQ-9 was 5 [38]. As this has not been established for the ACT and self-efficacy scales, we used the convention of 0.5 SD to define a MCID [39].

Results

Participants

Of the total patients screened, 17.3% (100/577) met eligibility criteria, consented to participate, and were randomized. The remaining 83% did not meet the inclusion criteria (62.3%, 297/477), declined to participate (4.0%, 19/477), dropped out before randomization (6.1%, 29/477), or did not have sufficient information to approach for recruitment (27.7%, 132/477; Figure 1). A total of 51 individuals were randomized to the control group and 49 to the intervention group. Table 1 presents the baseline characteristics for both the groups.

The distribution of characteristics between groups was similar; however, the proportion of individuals with a university degree and in the 0-39 age group was higher in the intervention group as compared with the control group (Table 1). Whereas the proportion of individuals who had one or more ED visits or hospitalizations for a pulmonary problem in 3 months before enrolment was similar between groups, in the 3 months before randomization, participants in usual care were more likely to overuse FABA in the 9 months before enrolment compared with those in the intervention group (12% vs 4%, respectively), although this difference was not statistically significant.

Figure 1. Consort diagram of participants in the My Asthma Portal (MAP) trial.

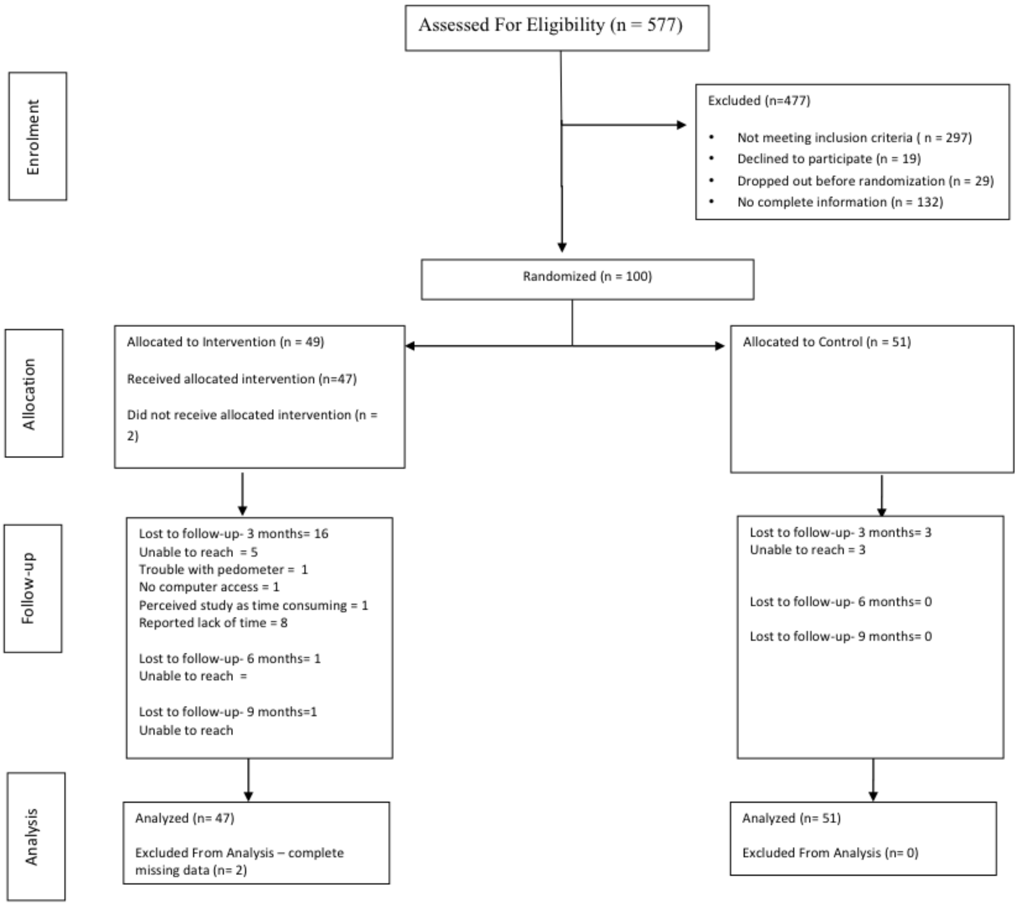


Table 1. Baseline characteristics of participants in the intervention and control groups.

Characteristic	Intervention (n=47) n (%)	Control (n=51) n (%)	*P* value
Sex			
Female	32 (68)	33 (65)	.72
Age group (years)			
0-39	20 (43)	14 (27)	.25
40-49	10 (21)	13 (25)	
50-59	12 (26)	15 (29)	
>60	5 (11)	9 (18)	
Smoking status			
Never smoker	26 (55)	26 (51)	.91
Current smoker	5 (11)	7 (14)	
Ex-smoker	15 (32)	16 (31)	
Missing	1 (2)	2 (4)	
Education			
>Higher school	32 (68)	26 (51)	.38
High school equivalent	13 (28)	22 (43)	
<High school	1 (2)	1 (2)	
Missing	1 (2)	2 (4)	
Self-reported health			
Excellent or very good	8 (17)	14 (27)	.31
Good	24 (51)	19 (37)	
Fair or poor	15 (32)	18 (35)	
Depression-PHQ[b]			
None (0-4)	22 (47)	27 (53)	.52
Mild (5-9)	19 (40)	13 (25)	
Moderate (10-14)	2 (4)	5 (10)	
Moderately severe or severe (15-19 or 20-7)	4 (8)	6 (12)	
Prescribed action plan for asthma at baseline			
Yes	33 (70)	41 (80)	.35
No	13 (28)	10 (20)	
Missing	1 (2)	0 (0)	
Number of asthma medications			
1	7 (15)	14 (27)	.25
2	19 (40)	17 (33)	
3	12 (26)	6 (12)	
≥4	6 (13)	10 (20)	
Missing	3 (6)	4 (8)	
Poorly controlled asthma			
No	38 (81)	36 (71)	.34
Yes	7 (15)	9 (18)	
Missing	2 (4)	6 (12)	

Characteristic	Intervention (n=47)	Control (n=51)	*P* value
	n (%)	n (%)	
ED[c] visits due to a respiratory problem			
0	37 (79)	38 (75)	.32
1	6 (13)	3 (6)	
≥2	2 (4)	4 (8)	
Missing	2 (4)	6 (12)	
FABA overuse[a]			
0	43 (91)	39 (76)	.48
1	2 (4)	6 (12)	
Missing	2 (4)	6 (12)	
Number of hospitalizations due to respiratory issues			
0	44 (94)	43 (84)	.42
1	1 (2)	1 (2)	
≥2	0 (0)	1 (2)	
Missing	2 (4)	6 (12)	

[a] PHQ: Patient Health Questionnaire.

[b] FABA (fast-acting bronchodilators) overuse: if the sum of the quantity for all FABAs dispensed was >750+ mg doses of salbutamol 100 mcg, 2 inhalations at a time, or the equivalent for other FABAs over a 9-month period.

[c] ED: emergency department.

Of the 100 individuals who were randomized, 16 were lost to follow-up at 3 months in the intervention group. Five individuals could no longer be reached, 1 reported having trouble with the pedometer and did not log in, 1 no longer had access to a computer, 1 reported that the study was time-consuming, and the remaining 8 individuals reported lack of time. In addition, one individual at 6 months and another at 9 months postrandomization could no longer be reached. Among these individuals, 5 did not log in to MAP at all, and the remainder logged in 1 to 36 times. Three individuals in the control group were lost to follow-up at 3 months postrandomization.

Process Evaluation

Access to MAP during the first 3 months ranged from 1 to 69 (mean 24, SD 16), and 0 to 38 between 3 and 6 months; among the 47 individuals who received the intervention, 16% had no logins between 3 and 6 months (Figure 2). Logins were more frequent in the first 4 weeks (mean 12, SD 8) and then tapered off thereafter (mean 36, SD 42). The average time spent logged into MAP was 7 minutes/login (range 20 seconds to 2 hours). The most frequently accessed components of the portal were the monitoring questionnaire and the asthma feedback center, with an average of 4.5 views in the first week, 2.2 in the second week, 1.5 from week 2 to week 7, and then 1 time per week for the remainder of the trial (Figure 3). The least frequently accessed component was the static Learning Center. In all instances when a monitoring questionnaire was started, it was completed. Among those who logged in, the proportion of participants who logged in at least once per week was the lowest in week 19 (56%) and the highest in week 1 (100%). Four individuals did not log in at all during the 6-month study period. Their characteristics were similar to those who used MAP except for depressive symptoms, which on average were lower (mean 5, SD 3) in those who did not log in compared with those who used MAP (mean 6, SD 5).

The most frequent reasons for initiating an email by the patient or the nurse was to discuss medications (30%), clarify the action plan (18%), monitor MAP feedback (15%), and conduct follow-ups related to control status (5%). About 61% of alerts to the nurse were related to medication adherence, 17% indicated a need to review the action plan, 11% were because the patient did not log in, and 7% were related to delayed initiation of the action plan by patients.

Usefulness ratings on the TAM were essentially unchanged at baseline and 6 months. A similar pattern was found for the response to the item “I intend to use My Asthma Web portal weekly to help manage my Asthma.”

Figure 2. Average number of logins by week over the 6-month study period.

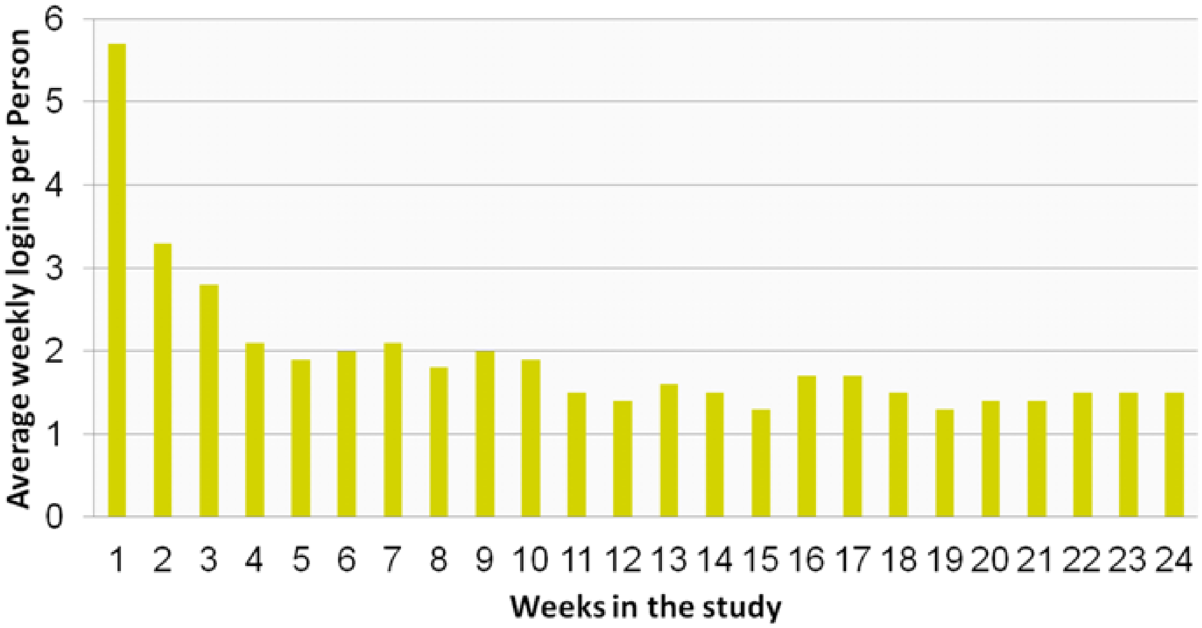
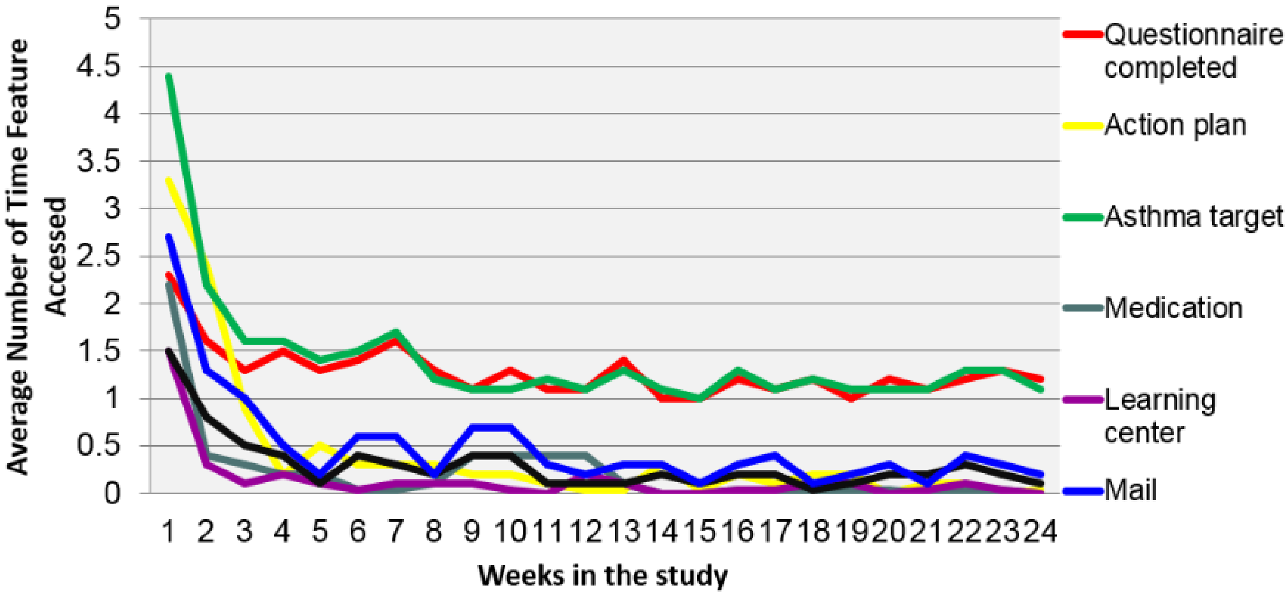


Figure 3. Usage of My Asthma Portal (MAP) features over time.



Impact of MAP on Asthma Control Using Administrative Data and Health-Related Quality of Life

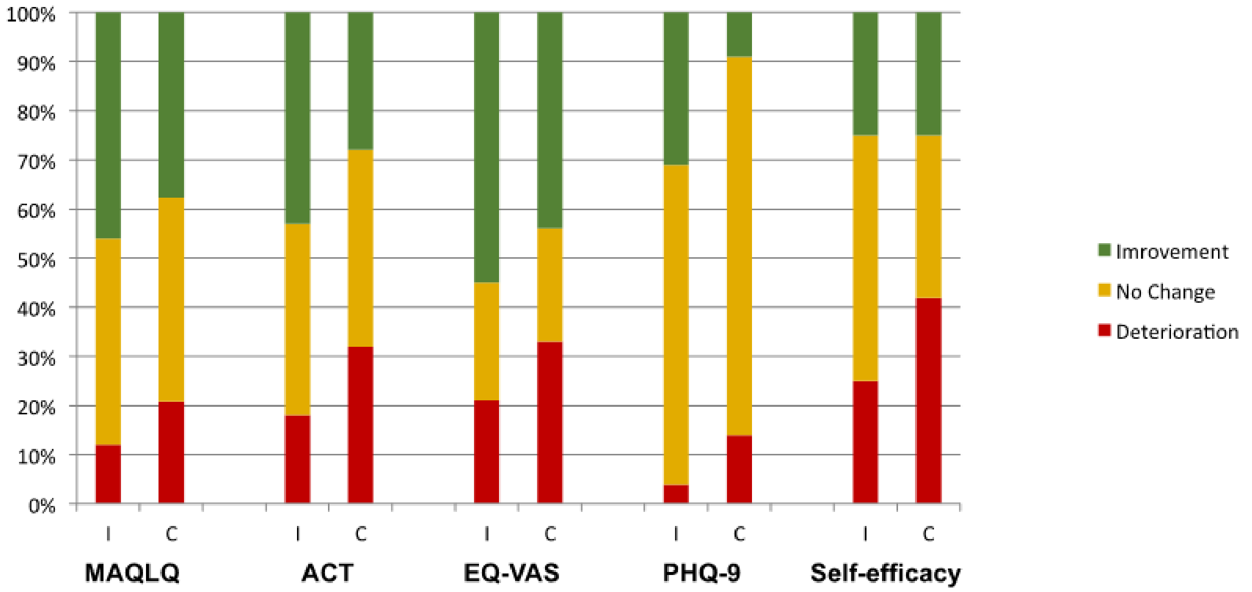
Multimedia Appendix 2 presents the scores for asthma control status and health-related quality of life (MAQLQ) over time by group. After adjustment for sex, age, asthma medications, education, and self-reported health, a significant improvement in MAQLQ score was evident between baseline and 3 months for the intervention group (mean change 0.67, 95% CI 0.36 to 0.98). Differences were not statistically significant for the intervention group between 3 and 6 months (mean change –0.01, 95% CI –0.35 to 0.32) or 6 and 9 months (mean change –0.12, 95% CI –0.46 to 0.22). There were no statistically significant differences over time for the control group. Further, there were no significant differences between groups over time for both MAQLQ and poor control status.

When models were reanalyzed with imputed data, the difference in MAQLQ scores between baseline and 3 months for the intervention group was still statistically significant after adjustment for sex, age, asthma medications, education, and self-reported health (mean change 0.52, 95% CI 0.20 to 0.84). In addition, a trend of decreasing scores between the 6 and 9 months was found (mean change –0.35, 95% CI –0.66 to –0.04) for the intervention group. No other statistically significant differences were found between time periods or between groups for MAQLQ scores. There was no significant effect of group, time, or group by time adjusted for sex, age, asthma medications, education, and self-reported health found for poor control status.

The results for differences in MAQLQ scores over time and between groups were similar with adjustment for depression (PHQ), self-efficacy, beliefs about asthma medication (BMQ), self-reported asthma control (ACT), and poor control status at baseline. In addition, depression (PHQ) (mean change –0.27, 95% CI –0.37 to –0.18 for a change of 5 units), self-efficacy (mean change 0.24, 95% CI 0.16 to 0.32), and self-reported asthma control symptoms (ACT) (mean change –0.25, 95% CI –0.30 to –0.20) were all significantly associated with asthma quality of life (MAQLQ). No significant effect of group, time, or group by time or for the explanatory variables was found for poor control status.

Figure 4 shows the proportion of individuals who showed a clinically meaningful improvement, no improvement, or deterioration on the patient-reported outcomes over the 6-month intervention period by group. For all measures, the intervention group had a higher proportion of individuals demonstrating a minimally clinically important difference compared with the control group. However, only the difference in depression (PHQ) at 6 months was statistically significant.

Figure 4. Minimal clinically important difference at 6 months on the patient reported outcomes. MAQLQ: Mini-Asthma Quality of Life Questionnaire; ACT: Asthma Control Test; EQ-VAS: EuroQol visual analogue scale; PHQ: Patient Health Questionnaire.



Discussion

Principal Findings

Our study found significant improvements in asthma quality of life over the initial 3 months among individuals using a Web-based asthma self-management intervention compared with those in the control group. There was also a higher proportion of individuals who had an improvement in depression, self-efficacy, and self-reported asthma control over time in the intervention group compared with the control group, suggesting that these were important mediating effects of the MAP intervention on improving asthma quality of life. However, there was no significant effect of the intervention on asthma control over time or for differences between groups in asthma control and asthma-related quality of life.

The proportion of poor control status seemed to worsen slightly in the intervention group and to improve in the usual care group during the study. This was not surprising, as the control group might also benefit from being included in a study. This suggested, however, that when patients reduced their interaction with MAP it might not be because they had learned how to self-manage their condition. Other potential reasons might be a loss of interest in the system, or feeling overwhelmed and unable to log in when symptoms worsened. Identifying mechanisms to keep individuals engaged and motivated to use systems such as MAP, which are relatively simple to integrate into daily activities, is important. Strategies found to be effective to increase the use of Web-based systems include primary task support that includes delivering the intervention in a stepwise approach as individuals' knowledge and behavioral changes progress and matching content and features of the intervention to the individuals' needs and health [16].

Ongoing access by participants using MAP supports the feasibility of using the system to deliver self-management for individuals who accept to start using the system. Also, the fact that logins decreased over time may reflect that once individuals learned about their medications and action plan they found less benefits of logging in to report symptoms and receive feedback that they became more familiar with over time. This reflects the importance of tailoring content and features of Web-based interventions over time to changes in individuals' knowledge, confidence, and symptoms as supported by previous reviews of self-management technology-based interventions [12,40].

Similar to our findings, a recent review of systematic reviews [12,40] showed that the 2 of the 7 studies that were conducted in an adult population [13,14] favored the Web-based intervention on asthma-related quality of life and on self-reported symptoms and asthma control. Previous work has shown that self-reported asthma control often overestimates the effect of interventions [41]. Our study was the first to examine the impact of a Web-based self-management program on asthma control independent of participant reporting [12]. We used a composite measure of control based on administrative data and an algorithm for overuse of rescue medication and ED use. While none of the previous studies examined the overuse of rescue medication as an outcome, they have evaluated the impact on ED visits. Similar to our study, Rasmussen et al [14] found no difference in ED visits between the intervention and control groups. Unlike our investigation that showed an association between self-efficacy, depression, and beliefs about medications with asthma-related quality of life, none of the previous studies to our knowledge examined these mediating effects as possible mechanisms by which Web-based self-management interventions enhance outcomes [12]. This also reflects the limited consideration of theory defining the core mechanisms by which technology-based interventions impact outcomes and guide the evaluation of self-management interventions [42]. Al-Durra et al (2015) found that only 20% (17/85) studies of Web-based asthma self-management interventions applied at least one model, framework, or construct of a behavioral change theory [43].

Furthermore, a consistent finding across systematic reviews of technology-enabled self-management is that interventions with multiple behavioral change techniques appear, on the whole, to be more effective than those using fewer techniques [12,40]. We also identified this as important during the iterative design process of MAP and in consultation with behavioral change experts on our team and incorporated several techniques within MAP, including self-monitoring, action planning, feedback in relation to symptoms and goals, environmental triggers, motivational messaging, and reward. Future work will focus on tailoring the use of behavioral change techniques and gradually combining techniques in response to individuals' clinical profile and preferences to help participants identify a sequence of behaviors to target. Tailoring the system to the needs of individuals as they take control of their health may help optimize individuals' perceived benefit of self-monitoring and staying connected to the care team as needed. In turn, this may increase the ongoing participant use of MAP and an opportunity to observe the long-term effects on asthma control.

Strengths and Limitations

Linking the mechanisms by which self-management interventions are expected to have an effect is important to guide the selection of functionalities that should be included in technology-based interventions. The features and decision support tools in MAP are based on the principles of self-management skill development [44], self-efficacy and motivational theory [45], and adult learning principles [44], and were integrated into clinical care with ongoing monitoring by a nurse case manager. These theories and the clinical process for asthma management guided the behavioral change techniques within MAP, the algorithm behind the frequency and nature of alerts to patients, and the interaction with the case manager [15]. While previous systems have included Web-based monitoring, most do not provide a comprehensive evaluation of asthma-related behaviors that can then be used to deliver automated alerts and decision support to patients and the care team [12]. In our study, we targeted asthma symptom monitoring and action plan use, but also included physical activity monitoring, which has been shown to be linked to averting exacerbations in the long term [46].

One of the limitations of this trial was that we did not evaluate adherence to asthma treatment. This limited our ability to evaluate the impact of MAP on adherence as an intermediate effect to improve asthma control. Also, the number of participants lost to follow-up and missing data resulted in a relatively small, but sufficient, sample size. Comparison of baseline characteristics and imputation to evaluate the impact of missing data provided insight into the characteristics of individuals who dropped out, and the implications for feasibility and the need to tailor the intervention. Previous reviews of computer-based asthma interventions reported attrition rates of up to 23% [12]. This is in keeping with other systematic reviews of computer-based interventions in other areas [12,40]. Our study had a higher attrition of 37%, which is still lower than the upper range of attrition found in nontechnology-based self-management interventions, which ranges from 0 to 54% [47,48]. Nonetheless, 16 participants were lost to follow-up in the intervention group compared with only 3 patients lost to follow-up in the control group. This finding implied that dropout attrition rate was more than 5 times higher in the intervention group compared with the control group. Eight individuals among those who dropped out reported a lack of time as a reason, indicating that the use of the system needs to be simplified to optimize use as part of daily activities. The remaining 7 individuals, who could not be reached and found the study time-consuming, might not have perceived MAP as beneficial. Initial perceived benefit could be used as a way to screen for potential users of MAP and identify ways of optimizing the fit of MAP to the needs of individuals with asthma. Only one person identified no access to a computer as a reason for dropout, indicating that, for some, lack of access to the Internet remains a potential barrier. Further, participants in this trial might have been those who were more likely to accept the use of technology, which might have limited the generalizability of the results to other patients with chronic conditions.

Another consideration is that care in specialty clinics was more intensive in terms of providing patients with case management and self-management support compared with primary care, which likely contributed to making it more challenging to find an effect. We would expect a larger difference in outcome between the Web-based tool and usual care in a primary care setting. Also, given the relatively small sample for this trial, about half had disease for >10 years and therefore had a greater opportunity to learn and improve asthma self-management behaviors prior to starting the trial. MAP may have greater impact among individuals in relatively earlier phases post-asthma diagnosis when learning self-management skills is new for them. This represents another promising avenue for future research.

Furthermore, the material and recommendations presented to participants in response to monitoring information were static for the entire 6-month period. Adapting the monitoring protocol and material and recommendations based on individuals' progression and time using the system may enhance usage and impact on outcomes.

Finally, from the time the trial was registered there were changes made to the design prior to the start of the trial, including expanding the inclusion for age from ≤60 to ≤69 and removing the criteria for full healthcare coverage and the need for the primary care physician to be using the asthma decision support.

Conclusions

There is growing interest in the potential of the Web-based and other digital media as a platform to deliver more tailored self-management support, while maintaining cost-effectiveness, with greater scope for integration into the everyday lives of those with asthma. The MAP intervention had a significant impact on improving asthma-related quality of life and other related outcomes including self-efficacy, depressive symptoms, and beliefs about medication. Furthermore, the effects of MAP were maintained at 6 months and began to decrease once the use of MAP ended between 6 and 9 months, suggesting that ongoing support might result in reduced exacerbations in the long term as opposed to usual care alone. Technology renders self-management support between clinical visits feasible. Future work is needed to identify tailored protocols for monitoring symptoms and health behaviors as individuals learn how to self-manage, and to assess the impact of MAP on asthma control and health-related quality of life beyond 6 months. Based on these results, a larger-cluster randomized trial will be designed with an updated version of MAP that will allow us to evaluate the long-term impact of tailored self-management, adjusting for multiple covariates over time.

Acknowledgments

This study was supported by a Canadian Institutes of Health Research operating grant (CIHR 89859, 217090, 204847), a Career Award for SA from the Fond de Recherche en Santé du Quebec (FRQS #13870), and a Canadian Foundation of Innovation Medical Informatics Infrastructure Grant (15649). The authors wish to acknowledge Pascale Rioux, Danielle Beaucage, and Dr Ron Olvenstein for their helpful input in the implementation of MAP and study recruitment.

Conflicts of Interest

None declared.

Multimedia Appendix 1

My Asthma Portal (MAP) monitoring business rules.

[PDF File (Adobe PDF File), 20KB - jmir_v18i12e313_app1.pdf]

Multimedia Appendix 2

Changes in asthma quality of life and asthma control status over time.

[PDF File (Adobe PDF File), 38KB - jmir_v18i12e313_app2.pdf]

References

1. Boulet LP, Bai TR, Becker A, Bérubé D, Beveridge R, Bowie DM, et al. What is new since the last (1999) Canadian Asthma Consensus Guidelines? Can Respir J 2001;8(Suppl A):5A-27A. [Medline: 11360044]
2. Ernst E. Complementary therapies for asthma: what patients use. J Asthma 1998;35(8):667-671. [Medline: 9860087]
3. Turner MO, Taylor D, Bennett R, Fitzgerald JM. A randomized trial comparing peak expiratory flow and symptom self-management plans for patients with asthma attending a primary care clinic. Am J Respir Crit Care Med 1998 Feb;157(2):540-546. [doi: 10.1164/ajrccm.157.2.9703060] [Medline: 9476870]
4. Tattersfield AE, Postma DS, Barnes PJ, Svensson K, Bauer CA, O'Byrne PM, et al. Exacerbations of asthma: a descriptive study of 425 severe exacerbations. The FACET International Study Group. Am J Respir Crit Care Med 1999 Aug;160(2):594-599. [doi: 10.1164/ajrccm.160.2.9811100] [Medline: 10430734]
5. Bonner S, Zimmerman BJ, Evans D, Irigoyen M, Resnick D, Mellins RB. An individualized intervention to improve asthma management among urban Latino and African-American families. J Asthma 2002 Apr;39(2):167-179. [Medline: 11990232]
6. Gibson PG. Monitoring the patient with asthma: an evidence-based approach. J Allergy Clin Immunol 2000 Jul;106(1 Pt 1):17-26. [doi: 10.1067/mai.2000.108307] [Medline: 10887300]
7. Hartert TV, Windom HH, Peebles RS, Freidhoff LR, Togias A. Inadequate outpatient medical therapy for patients with asthma admitted to two urban hospitals. Am J Med 1996 Apr;100(4):386-394. [Medline: 8610724]
8. Crain EF, Kercsmar C, Weiss KB, Mitchell H, Lynn H. Reported difficulties in access to quality care for children with asthma in the inner city. Arch Pediatr Adolesc Med 1998 Apr;152(4):333-339. [Medline: 9559707]
9. Dinkevich EI, Cunningham SJ, Crain EF. Parental perceptions of access to care and quality of care for inner-city children with asthma. J Asthma 1998;35(1):63-71. [Medline: 9513584]

10. FitzGerald JM, Boulet L, McIvor RA, Zimmerman S, Chapman KR. Asthma control in Canada remains suboptimal: the Reality of Asthma Control (TRAC) study. Can Respir J 2006;13(5):253-259 [FREE Full text] [Medline: 16896426]
11. Eccles MP, Armstrong D, Baker R, Cleary K, Davies H, Davies S, et al. An implementation research agenda. Implement Sci 2009 Apr 07;4:18 [FREE Full text] [doi: 10.1186/1748-5908-4-18] [Medline: 19351400]
12. Morrison D, Wyke S, Agur K, Cameron EJ, Docking RI, Mackenzie AM, et al. Digital asthma self-management interventions: a systematic review. J Med Internet Res 2014;16(2):e51 [FREE Full text] [doi: 10.2196/jmir.2814] [Medline: 24550161]
13. van der Meer V, Bakker MJ, van den Hout WB, Rabe KF, Sterk PJ, Kievit J, SMASHING (Self-Management in Asthma Supported by Hospitals, ICT, NursesGeneral Practitioners) Study Group. Internet-based self-management plus education compared with usual care in asthma: a randomized trial. Ann Intern Med 2009 Jul 21;151(2):110-120. [Medline: 19620163]
14. Rasmussen LM, Phanareth K, Nolte H, Backer V. Internet-based monitoring of asthma: a long-term, randomized clinical study of 300 asthmatic subjects. J Allergy Clin Immunol 2005 Jun;115(6):1137-1142. [doi: 10.1016/j.jaci.2005.03.030] [Medline: 15940125]
15. Ahmed S, Bartlett SJ, Ernst P, Paré G, Kanter M, Perreault R, et al. Effect of a web-based chronic disease management system on asthma control and health-related quality of life: study protocol for a randomized controlled trial. Trials 2011;12:260 [FREE Full text] [doi: 10.1186/1745-6215-12-260] [Medline: 22168530]
16. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152. [doi: 10.2196/jmir.2104] [Medline: 23151820]
17. Juniper EF, Price DB, Stampone PA, Creemers JP, Mol SJ, Fireman P. Clinically important improvements in asthma-specific quality of life, but no difference in conventional clinical indexes in patients changed from conventional beclomethasone dipropionate to approximately half the dose of extrafine beclomethasone dipropionate. Chest 2002 Jun;121(6):1824-1832. [Medline: 12065345]
18. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
19. Hendrickson AR, Massey PD, Cronan TP. On the test-retest reliability of perceived usefulness and perceived ease of use scales. MIS Quarterly 1993 Jun;17(2):227-230. [doi: 10.2307/249803]
20. Adams DA, Nelson RR, Todd PA. Perceived usefulness, ease of use, and usage of information technology: a replication. MIS Quarterly 1992 Jun;16(2):227-247. [doi: 10.2307/249577]
21. Szajna B. Empirical evaluation of the revised technology acceptance model. Management Science. 1996 Jan 1996;42(1):85. [doi: 10.1287/mnsc.42.1.85]
22. Spitzer WO, Suissa S, Ernst P, Horwitz RI, Habbick B, Cockcroft D, et al. The use of beta-agonists and the risk of death and near death from asthma. N Engl J Med 1992 Feb 20;326(8):501-506. [doi: 10.1056/NEJM199202203260801] [Medline: 1346340]
23. Juniper EF, Guyatt GH, Cox FM, Ferrie PJ, King DR. Development and validation of the Mini Asthma Quality of Life Questionnaire. Eur Respir J 1999 Jul;14(1):32-38 [FREE Full text] [Medline: 10489826]
24. Ehrs PO, Nokela M, Ställberg B, Hjemdahl P, Wikström JE. Brief questionnaires for patient-reported outcomes in asthma: validation and usefulness in a primary care setting. Chest 2006 Apr;129(4):925-932. [doi: 10.1378/chest.129.4.925] [Medline: 16608940]
25. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum 1989 Jan;32(1):37-44. [Medline: 2912463]
26. Tobin DL, Wigal JK, Winder JA, Holroyd KA, Creer TL. The “asthma self-efficacy scale”. Ann Allergy 1987 Oct;59(4):273-277. [Medline: 3662130]
27. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol 2004 Jan;113(1):59-65. [doi: 10.1016/j.jaci.2003.09.008] [Medline: 14713908]
28. Schatz M, Sorkness CA, Li JT, Marcus P, Murray JJ, Nathan RA, et al. Asthma Control Test: reliability, validity, and responsiveness in patients not previously followed by asthma specialists. J Allergy Clin Immunol 2006 Mar;117(3):549-556. [doi: 10.1016/j.jaci.2006.01.011] [Medline: 16522452]
29. Juniper EF. Using humanistic health outcomes data in asthma. Pharmacoeconomics 2001;19(Suppl 2):13-19. [Medline: 11700784]
30. Emilsson M, Berndtsson I, Lötvall J, Millqvist E, Lundgren J, Johansson A, et al. The influence of personality traits and beliefs about medicines on adherence to asthma treatment. Prim Care Respir J 2011 Jun;20(2):141-147 [FREE Full text] [doi: 10.4104/pcrj.2011.00005] [Medline: 21311839]
31. 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]
32. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 2012 Feb 21;184(3):E191-E196 [FREE Full text] [doi: 10.1503/cmaj.110829] [Medline: 22184363]

33. Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(96)00822-6] [Medline: 10158943]
34. Wilchesky M, Tamblyn RM, Huang A. Validation of diagnostic codes within medical services claims. J Clin Epidemiol 2004 Feb;57(2):131-141. [doi: 10.1016/S0895-4356(03)00246-4] [Medline: 15125622]
35. Regie de l'assurance-maladie du Quebec. Statistiques Annuelles. Regie de l'assurance-maladie du Quebec 2000:46-48.
36. Jiang J. Linear and generalized linear mixed models and their applications. New York: Springer; 2007.
37. Schafer JL. Analysis of incomplete multivariate data. Boca Raton, Florida: CRC press; 1997.
38. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Med Care 2004 Dec;42(12):1194-1201. [Medline: 15550799]
39. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003 May;41(5):582-592. [doi: 10.1097/01.MLR.0000062554.74615.4C] [Medline: 12719681]
40. McDermott MS, While AE. Maximizing the healthcare environment: a systematic review exploring the potential of computer technology to promote self-management of chronic illness in healthcare settings. Patient Educ Couns 2013 Jul;92(1):13-22. [doi: 10.1016/j.pec.2013.02.014] [Medline: 23566427]
41. Fayers P, Machin D. Quality of life: the assessment, analysis and interpretation of patient-reported outcomes. England: John Wiley & Sons; 2013.
42. Vassilev I, Rowsell A, Pope C, Kennedy A, O'Cathain A, Salisbury C, et al. Assessing the implementability of telehealth interventions for self-management support: a realist review. Implement Sci 2015;10:59 [FREE Full text] [doi: 10.1186/s13012-015-0238-9] [Medline: 25906822]
43. Al-Durra M, Torio MB, Cafazzo JA. The use of behavior change theory in Internet-based asthma self-management interventions: a systematic review. J Med Internet Res 2015 Apr 02;17(4):e89 [FREE Full text] [doi: 10.2196/jmir.4110] [Medline: 25835564]
44. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [Medline: 12401421]
45. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [Medline: 12867348]
46. Bacon SL, Lemiere C, Moullec G, Ninot G, Pepin V, Lavoie KL. Association between patterns of leisure time physical activity and asthma control in adult patients. BMJ Open Resp Res 2015 Jul 24;2(1):e000083. [doi: 10.1136/bmjresp-2015-000083] [Medline: 26244098]
47. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013;3:CD008776. [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
48. Wieland LS, Falzon L, Sciamanna CN, Trudeau KJ, Brodney S, Schwartz JE, et al. Interactive computer-based interventions for weight loss or weight maintenance in overweight or obese people. Cochrane Database Syst Rev 2012;8:CD007675 [FREE Full text] [doi: 10.1002/14651858.CD007675.pub2] [Medline: 22895964]

Abbreviations

ACT: Asthma Control Test
ED: emergency department
EQ-VAS: EuroQol visual analogue scale
MAP: My Asthma Portal
MAQLQ: Mini-Asthma Quality of Life Questionnaire
PHQ: Patient Health Questionnaire

Edited by G Eysenbach; submitted 16.05.16; peer-reviewed by M Al-Durra, S Omboni; comments to author 30.06.16; revised version received 11.10.16; accepted 29.10.16; published 01.12.16.

Please cite as:
Ahmed S, Ernst P, Bartlett SJ, Valois MF, Zaihra T, Paré G, Grad R, Eilayyan O, Perreault R, Tamblyn R
The Effectiveness of Web-Based Asthma Self-Management System, My Asthma Portal (MAP): A Pilot Randomized Controlled Trial
J Med Internet Res 2016;18(12):e313
URL: http://www.jmir.org/2016/12/e313/
doi:10.2196/jmir.5866
PMID:27908846

Original Paper

An eHealth Project on Invasive Pneumococcal Disease: Comprehensive Evaluation of a Promotional Campaign

Donatella Panatto[1*], PhD; Alexander Domnich[1*], MD; Roberto Gasparini[1], MD; Paolo Bonanni[2], MD; Giancarlo Icardi[1], MD; Daniela Amicizia[1], MD; Lucia Arata[1], MD; Stefano Carozzo[1], MD; Alessio Signori[1], PhD; Angela Bechini[2], PhD; Sara Boccalini[2], PhD

[1]Department of Health Sciences, University of Genoa, Genoa, Italy

[2]Department of Health Sciences, University of Florence, Florence, Italy

[*]these authors contributed equally

Corresponding Author:
Donatella Panatto, PhD
Department of Health Sciences
University of Genoa
Via Pastore, 1
Genoa, 16132
Italy
Phone: 39 0103538109
Fax: 39 0103538541
Email: panatto@unige.it

Abstract

Background: The recently launched *Pneumo Rischio* eHealth project, which consists of an app, a website, and social networking activity, is aimed at increasing public awareness of invasive pneumococcal disease (IPD). The launch of this project was prompted by the inadequate awareness of IPD among both laypeople and health care workers, the heavy socioeconomic burden of IPD, and the far from optimal vaccination coverage in Italy, despite the availability of safe and effective vaccines.

Objective: The objectives of our study were to analyze trends in *Pneumo Rischio* usage before and after a promotional campaign, to characterize its end users, and to assess its user-rated quality.

Methods: At 7 months after launching *Pneumo Rischio*, we established a 4-month marketing campaign to promote the project. This intervention used various approaches and channels, including both traditional and digital marketing strategies. To highlight usage trends, we used different techniques of time series analysis and modeling, including a modified Mann-Kendall test, change-point detection, and segmented negative binomial regression of interrupted time series. Users were characterized in terms of demographics and IPD risk categories. Customer-rated quality was evaluated by means of a standardized tool in a sample of app users.

Results: Over 1 year, the app was accessed by 9295 users and the website was accessed by 143,993 users, while the project's Facebook page had 1216 fans. The promotional intervention was highly effective in increasing the daily number of users. In particular, the Mann-Kendall trend test revealed a significant ($P \leq .01$) increasing trend in both app and website users, while change-point detection analysis showed that the first significant change corresponded to the start of the promotional campaign. Regression analysis showed a significant immediate effect of the intervention, with a mean increase in daily numbers of users of 1562% (95% CI 456%-4870%) for the app and 620% (95% CI 176%-1777%) for the website. Similarly, the postintervention daily trend in the number of users was positive, with a relative increase of 0.9% (95% CI 0.0%-1.8%) for the app and 1.4% (95% CI 0.7%-2.1%) for the website. Demographics differed between app and website users and Facebook fans. A total of 69.15% (10,793/15,608) of users could be defined as being at risk of IPD, while 4729 users expressed intentions to ask their doctor for further information on IPD. The mean app quality score assigned by end users was approximately 79.5% (397/500).

Conclusions: Despite its specific topic, *Pneumo Rischio* was accessed by a considerable number of users, who ranked it as a high-quality project. In order to reach their target populations, however, such projects should be promoted.

(J Med Internet Res 2016;18(12):e316) doi:10.2196/jmir.6205

KEYWORDS

invasive pneumococcal disease; pneumococcus; eHealth; mHealth; mobile app

Introduction

Invasive Pneumococcal Disease: High Burden and Low Awareness

Streptococcus pneumoniae, also known as pneumococcus, is an important human pathogen. It can cause both noninvasive (eg, otitis media, sinusitis, pneumonia) and invasive pneumococcal disease (IPD), which is described as the presence of pneumococcus in normally sterile body fluids. IPD has a variety of clinical presentations, the most common being meningitis, bacteremia, and bacteremic pneumonia [1,2]. It has been estimated [3] that approximately 1.6 million people worldwide die of IPD each year. In Italy, IPD is the invasive bacterial infection that carries the highest burden in terms of morbidity. Indeed, in 2014, a total of 952 IPD cases were notified (accounting for 78% of all cases), while cases of invasive diseases caused by *Neisseria meningitidis* accounted for 163 (13%) and those caused by *Haemophilus influenzae* accounted for 105 (9%) [4]. The risk of developing IPD is unevenly distributed among different population groups, being significantly higher among young children, the elderly, and people with several underlying medical conditions and health-compromising behaviors [5,6].

Vaccination is the only public health measure able to drastically reduce the incidence of IPD [7,8] and is highly recommended [9] for the above-mentioned at-risk population groups. However, immunization rates remain relatively low in both Italy [9] and other developed countries, including the United States [10]. The reasons for this are very probably multiple and of different nature, although inadequate knowledge and awareness of IPD among both health care practitioners and patients seems to be a major factor. Indeed, Lode et al [11] found that the main obstacles to vaccination among laypeople were scant awareness of vaccine availability, insufficient IPD risk perception, and lack of recommendation by general practitioners (GPs). The same research group [11] reported scant awareness of terms for pneumococcal diseases, with only 50% of GPs knowing the term IPD. Nichol et al [12] found that recommendation by a health care provider enhanced pneumococcal vaccine adherence among adults at risk of IPD. A systematic review of the determinants of pneumococcal vaccination [13] confirmed this finding, revealing that strong recommendation by GPs is an effective strategy for increasing immunization rates.

Pneumo Rischio Project

We have previously shown [14] that, despite the heavy health and socioeconomic burden of IPD in Italy [4], considerably less information is available on this disease than on other infectious pathologies with lower incidence rates. In order to fill this information gap—that is, the discrepancy between disease occurrence and available information [15]—we recently developed and launched an eHealth project called *Pneumo Rischio* [14]. Implementation of this project was also prompted by the above-described scant awareness of IPD among both laypeople and health care professionals [11-13] and the low vaccination coverage in Italy [9], despite the availability of safe and effective vaccines.

Pneumo Rischio was launched on February 26, 2015 in the three main app stores available in Italy, namely Google Play (Google Inc, Mountain View, CA, USA) [16], iTunes (Apple Inc, Cupertino, CA, USA) [17], and Microsoft (Microsoft Corporation, Redmond, WA, USA) [18] stores. The development process and main features of *Pneumo Rischio* have been reported elsewhere [14]. Briefly, the app was conceived to be maximally functional and easy to use and navigate. The core component of the app is a checker, which is designed to estimate the personal risk of contracting IPD and inform its customers in a user-friendly manner of communication. Once users have completed the checker, they can send the complete output to their doctor by email. The ultimate goal of the app is to increase community awareness of IPD. We created a sister website [19] to ensure a higher population coverage [14]. Moreover, we have also set up a Facebook account (*Pneumo Rischio* product page, subcategory app page) [20] to share IPD-related information and resources. Figure 1 shows screenshots of the *Pneumo Rischio* app, website, and Facebook page.

Figure 1. *Pneumo Rischio* app (A), website (B), and Facebook page (C).

Questa applicazione è facile da usare e con pochi click permette di

calcolare	sensibilizzare	stimolare
il rischio potenziale di contrarre una malattia invasiva da pneumococco (es. polmoniti batteriemiche, meningiti e sepsi)	sul rischio di contrarre le malattie da pneumococco	a chiedere ulteriori informazioni al proprio medico di riferimento in merito alle malattie da pneumococco e alla loro prevenzione

Rationale and Aims

To be effective, an app concerning health and health care should not only be technically efficient but also (1) be evidence based [21], (2) satisfy different aspects of objective and subjective quality [22], (3) be useful to and targeted at its end users [23], and (4) be downloadable and cover as much of the population as possible [24]. Undoubtedly, the first requirement should be satisfied in the initial stages of development; app developers, for example, could involve experts in a given health care field, employ methods of evidence-based medicine to build the product content, and test their app for quality and perceived usefulness in the phase of testing app prototypes. With regard to *Pneumo Rischio*, which proved to be a user-friendly product, we tried to make it evidence based by involving IPD experts in drafting content and features of the project and by carrying out a comprehensive literature review [14]. However, reaching this goal may not guarantee the achievement of the other three above-mentioned requirements. Thus, quality and usefulness evaluations made by real-world users after the app has been launched may differ from assessments made in the prelaunch phase, owing to poor representativeness of the target population or the fact that the judgments of real-world users may be scattered and highly subjective [22]. Continuous monitoring of users' feedback and specific instruments may help assess user-perceived app quality and usefulness and may prompt modifications of the app content and functions in future updates. The fourth aspect (ie, the number of users) is of particular relevance to public health and preventive medicine.

In the light of the above considerations, we aimed to provide some potentially useful insights into assessing an app/eHealth project from different points of view. Specifically, with regard to the *Pneumo Rischio* project, our goals were (1) to describe and analyze the trend in usage and its determinants, (2) to characterize users and their representativeness, and (3) to evaluate users' ratings of the quality of the app.

Methods

Marketing Campaign to Increase Population Exposure to *Pneumo Rischio*

In the first 7 months after the project was launched, we had no defined marketing strategy for its promotion; we only occasionally presented it at conferences, and we made some Web press releases. From the fourth week of September 2015, we established a more intense 4-month marketing campaign (henceforth referred to as the intervention). This used various approaches and channels, including both traditional and digital marketing strategies. The traditional approach consisted of presenting the app and the first results of its use at meetings, congresses [25,26], and postgraduate courses, participating in an eHealth competition, and advertising at physicians' offices. The online component included online advertising (Google AdWords, Google AdSense, Google AdMob, Facebook Ads) and social media (primarily Facebook). While pay-per-click ads are a recognized and cost-effective means of advertising [27], social networking offers a unique opportunity for social promotion [28] and has proved to be a feasible recruitment option [29]. We used the *Pneumo Rischio* Facebook page to periodically post stylistically coherent messages concerning IPD risk factors. Most posts could be classified as designed questions [30] (eg, "Can my *Pneumo Rischio* increase if…?"). Moreover, we posted banners on some popular thematic and informational portals.

Usage: Trends and Determinants

We used Google Analytics data, which are rigorously anonymous and presented in aggregated form, to record the number of users, number of sessions, users' demographic characteristics, and the risk (related to IPD) profiles of *Pneumo Rischio* users. The main unit of analysis was the daily (n=365) number of app and website users. We considered app and website users separately, since they have different characteristics and usage patterns [14].

We exploited different techniques of time series analysis to highlight the usage trend. The modified Mann-Kendall test [31] for serially autocorrelated data was used to assess the statistical significance of the trend in the number of *Pneumo Rischio* users over time. We then performed change-point detection analysis to locate points at which the statistical properties (both mean and variance) of the time series changed [32]. For this purpose, we used the binary segmentation approach.

We subsequently carried out segmented negative binomial regression analysis (to account for overdispersion) of interrupted time series in order to quantify immediate and time-related changes in the counts of users after the launch of the intervention. Specifically, the equation was formulated as follows:

$U_t= \beta_0+ \beta_1\times \text{time}_t+ \beta_2\times \text{intervention}_t+ \beta_3\times \text{time postintervention}_t+ \varepsilon_t,$

where U_t is the mean number of app or website users on day *t*; *time* is the day from the *Pneumo Rischio* launch; *intervention* is a binary variable indicating time *t* before (0) or after (1) the start of the intervention; *postintervention* is the number of days after the start of the intervention on day *t*, expressed as 0 before the intervention and (time-detected change point) after the intervention; β_0 is the baseline level of users at time 0; β_1 is the day-by-day change in the mean number of users before the intervention; β_2 is the level change in the mean number of users immediately after the intervention; β_3 is the trend change in the day-by-day mean number of users after the intervention in comparison with the preintervention period; and ε is the error term [33]. Since residuals of both models were heteroskedastic and serially correlated, the inferential testing of model parameters used heteroskedasticity-autocorrelation consistent standard errors.

Subsequently, we described usage of the *Pneumo Rischio* Facebook product page; we collected these anonymous and aggregated data from Facebook Insights. We also investigated how different types of Facebook posts—ie, "Photos" (posts containing photos), "SharedVideo" (posts containing videos), "Notes" (posts with HTML capability and no word limit), and "Links" (posts with links to other sites)—can affect the engagement of visitors with the *Pneumo Rischio* Facebook page. Posts were also dichotomized by type into personalized (posts containing personal or possessive adjectives or pronouns; eg, "my," "mine") and neutral posts (without any personal determiners). Since the distribution of people's engagement with Facebook posts was highly skewed, we applied the nonparametric Mann-Whitney *U* test and computed the effect size as Cohen $r_c= z/\sqrt{n}$ [34].

Pneumo Rischio: End Users' Characteristics

Next, we characterized *Pneumo Rischio* users (app and website users and Facebook fans) in terms of sex, age, and IPD risk categories; that is, low (healthy adults), medium (healthy elderly and immunocompetent people of any age with chronic conditions), and high (immunocompromised people of any age) [35]. These data were collected from Google Analytics (app and website) and Facebook Insights (Facebook page). As a theoretical measure of the effectiveness of the project, *Pneumo Rischio* users were asked (once the final result had been visualized) whether they intended to ask their GP for further information on IPD and its prevention. The second proxy measure of project effectiveness was the number of emails with detailed IPD risk profiles sent. This proxy measure is based on the assumption that a user sends a personal health-related record to his or her GP.

Objective and Subjective Quality as Defined by End Users

The quality of the app was assessed by means of the Mobile Application Rating Scale (MARS) [22], the Italian version of which has recently been validated [36]. The scale consists of 23 Likert-type items on a 5-point range (from 1, “poor,” to 5, “excellent”) and assesses app quality in 4 objective dimensions (engagement, functionality, aesthetics, and information) and 1 subjective quality dimension. A summary score for each dimension is obtained by averaging the corresponding scores of single items. The MARS total score is obtained by averaging the summary scores on the 4 objective quality dimensions. The Italian version of MARS has shown good psychometric properties; the MARS total score has been seen to have an intraclass correlation coefficient of .96, Cronbach alpha of .90, and acceptable levels of convergent, divergent, discriminative, known-groups validity, and scalability [36].

Since the MARS was originally intended to be used by trained professionals, a simplified training-free version of the scale was also created in order to obtain app-user quality ratings [37]. The two MARS versions are very similar; however, the user version uses simpler wording, contains fewer technical terms, and omits 3 items on the information subscale (accuracy of the app description in the app store, goals, and evidence base) [22,37]. Moreover, both MARS versions have an app-specific section that is adjustable to research aims. In this study, we evaluated the potential impact of the app on users’ knowledge, attitudes, awareness and behavior. Considering the above similarities, we supposed that most psychometric properties of the professional version would be transferable to the user version.

Subsequently, we assessed the customer-determined quality of the app. For this purpose, we enrolled participants during conferences and courses (health care professionals) and lectures (students in various graduate and postgraduate courses). Enrolled participants were instructed to navigate in all app components and functions for at least 10 minutes, to examine the description/definition of the app quality provided for each subscale, and then to fill in an anonymous paper-and-pencil survey form. Before evaluating the app, participants were not aware of the *Pneumo Rischio* project.

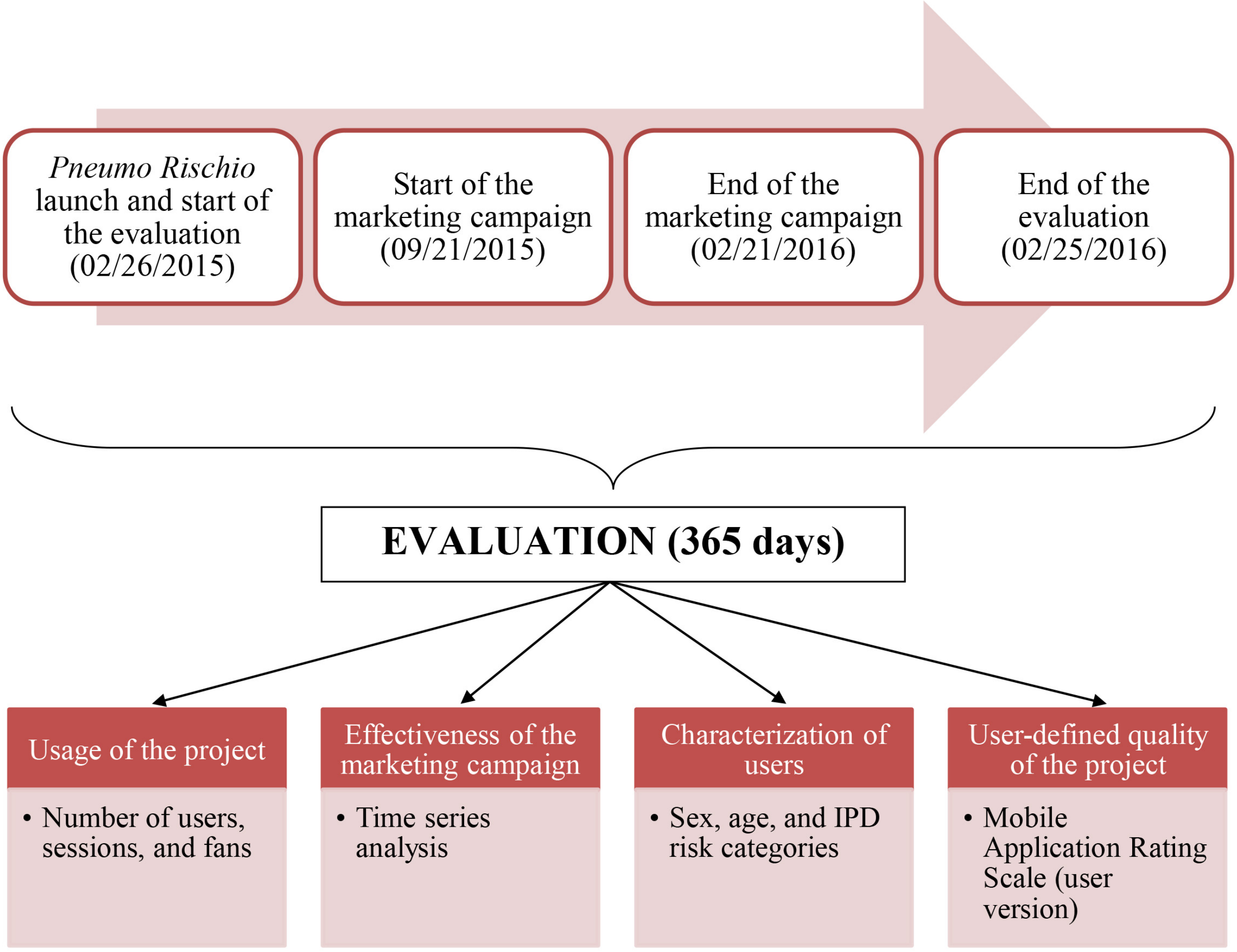
Participation in this nonbiomedical, noninterventional study was voluntary, and anonymity was guaranteed. Ethical approval for this study was not required, since it focused only on the quality evaluation of an existing service, which is freely available in the public domain, with no potential risks for participants. We collected no personal or sensitive data.

We express approximately normally distributed summary scores as means with standard deviations, and we summarized single Likert scale items as medians with interquartile ranges (IQRs). Average MARS summary scores exceeding 3.0 points (60%) were regarded as satisfactory [38]. We compared the MARS total scores between sexes and professions (health care vs non-health care) by means of t test, and calculated the Pearson r correlation coefficient between the score and the age of raters. The effect size for the t test was quantified by means of Cohen d. Finally, we constructed multivariable linear models (chosen by minimizing the corrected Akaike information criterion) with heteroskedasticity consistent standard errors to predict the MARS total and subscale scores.

All statistical analyses and modeling procedures were performed in R environment (R Foundation for Statistical Computing).

Figure 2 schematizes the chronology and assessment methods of the study.

Figure 2. Chronology and assessment methods of the study of usage of the *Pneumo Rischio* project. IPD: invasive pneumococcal disease.



Results

Usage: Trends and Determinants

On February 25, 2016, exactly 1 year after being launched, the app had been downloaded 9295 times and 10,090 sessions were run (1.09 sessions/user). On average, a session lasted 04:01 minutes. The website had been visited by 143,993 users, who ran 150,790 sessions (1.05 sessions/user), with a mean number of 2.13 pages per session. Website sessions lasted approximately four times less (01:02 minutes) than an average app session.

The modified Mann-Kendall trend test showed a statistically significant increasing trend in both app (corrected z=3.33, P<.001) and website (corrected z=2.75, P=.01) users. Change-point detection analysis revealed 5 significant changes in the daily time series of both app and website use. The first point corresponded to the start of the intervention (fourth week of September 2015). Analogously, segmented regression analysis confirmed the effectiveness of the intervention (Figure 3). As Table 1 shows, the baseline level of app and website users was statistically different from zero. The preintervention trends displayed opposite patterns, being negative for the app and positive for the website; these parameter estimates did not, however, reach an alpha <.05. Estimated usage grew significantly (P<.001) immediately after the start of the intervention: we estimated an increase in daily numbers users of 1562% (95% CI 456%-4870%) for the app and 620% (95% CI 176%-1777%) for the website. Moreover, the postintervention day-by-day trend in the number of users was also positive and statistically significant for the website, with a 1.4% (95% CI 0.7%-2.1%) increase, but not for the app, with only a 0.9% (95% CI 0.0%-1.8%) increase.

On February 25, 2016, the *Pneumo Rischio* Facebook page had 1216 likes registered. The daily number of new likes correlated highly with both app (r=.60, 95% CI .53-.66) and website users (r=.59, 95% CI .52-.65). Since the start of the intervention, 30 posts had been published; most of these were of the Notes (n=16) and Photos (n=12) types, while there was only 1 SharedVideo and 1 Link. We excluded the SharedVideo and Link types of posts from the analysis, owing to their singularity. We categorized 9 posts as personalized, and the remaining 19 as neutral. Photo posts engaged a significantly (P<.001) higher median number of users than Notes (1075 vs 376), and the effect size was large (r_c=0.74). By contrast, the higher number of users with personalized (median 631) rather than neutral (median 540) posts was not statistically significant (P=.76).

Figure 3. Daily observed and predicted numbers of *Pneumo Rischio* app and website users, February 26, 2015 to February 25, 2016.

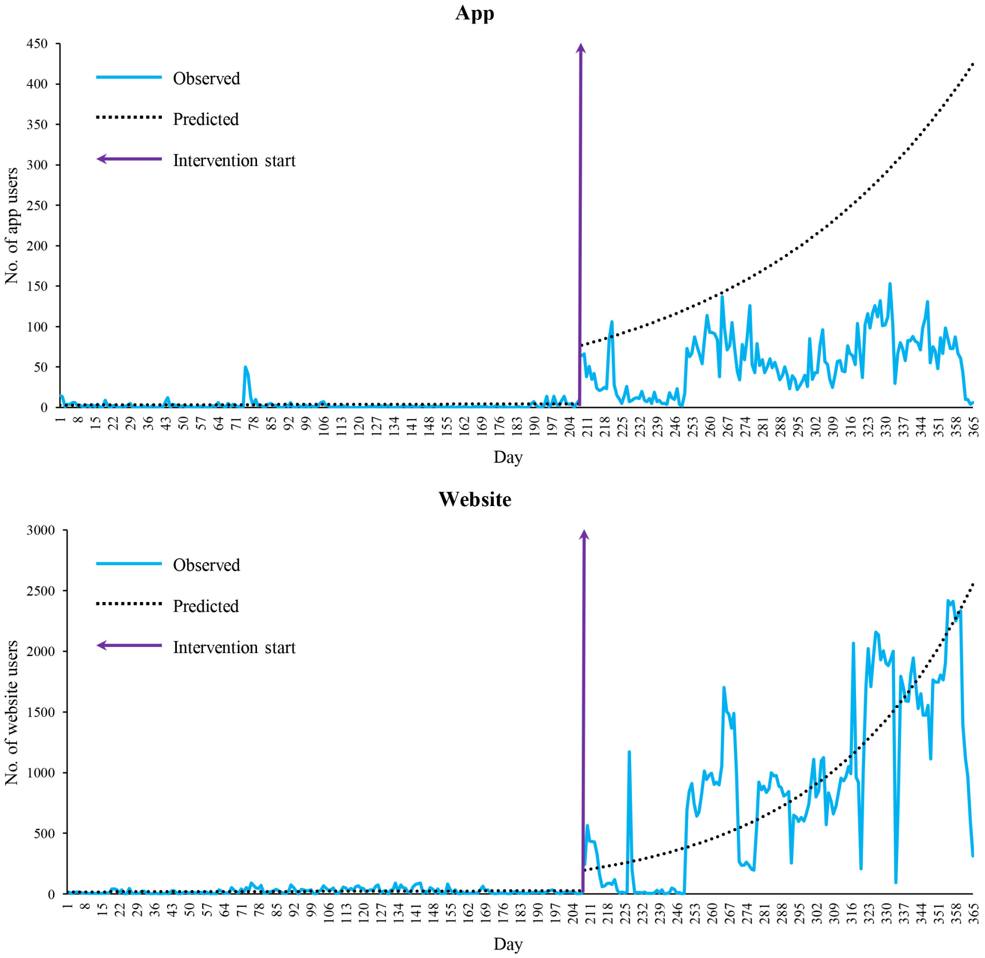


Table 1. Segmented regression analysis to predict the number of app and website users.

Parameter	App		Website	
	β (SE)	*z* (*P* value)	β (SE)	*z* (*P* value)
Baseline level	1.091 (0.449)	2.43 (.02)	2.804 (0.256)	10.95 (<.001)
Baseline trend	–0.002 (0.004)	–0.54 (.59)	0.002 (0.002)	1.10 (.27)
Postintervention level change	2.811 (0.559)	5.03 (<.001)	1.973 (0.489)	4.03 (<.001)
Postintervention trend change	0.009 (0.005)	1.86 (.06)	0.014 (0.004)	3.91 (<.001)

Pneumo Rischio: End Users' Characteristics

Demographic profiles differed between app users, website visitors, and Facebook fans (Table 2). In terms of the sex of users, women downloaded the app 1.5 times more frequently than men, and three-quarters of Facebook fans were female.

Conversely, a higher number of men visited the website. The most numerous age class of both app and website users was that of adults aged 25-34 years. *Pneumo Rischio* Facebook fans were older than *Pneumo Rischio* users, the most representative age class being 45-54 years, followed by the 55-64 and 35-44 age classes. It is encouraging that approximately 12% (145/1216) of the Facebook fans were aged ≥65 years.

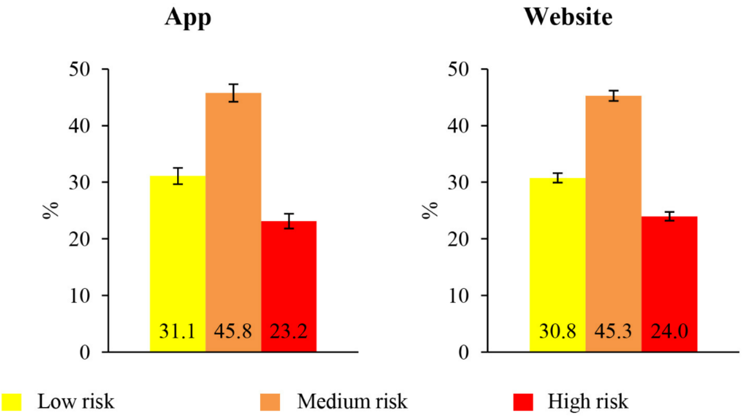
Approximately 70% (app: 2732/3965; website: 8061/11,643) of users may be defined as being at risk of IPD. The distribution of risk categories was similar between app and website users (Figure 4). A total of 2617 app users (25.9%, 95% CI 25.1%-26.8% of all sessions) and 2112 website users (1.4%, 95% CI 1.3%-1.5% of all sessions) stated that they would ask their GP for further information on IPD and its prevention. A total of 2142 (1700 for the app and 442 for the website) emails with detailed IPD profiles were sent.

Table 2. Age and sex distribution of *Pneumo Rischio* users.

Parameter	App users		Website users	Facebook fans	
	%	n/total	%[a]	%	n/total
Sex					
Male	39.2	166/423	54.1	24.84	302/1216
Female	60.8	257/423	45.9	74.92	911/1216
Unknown	–		–	0.25	3/1216
Age class, years					
18-24	7.0	29/412	27.5	3.13	38/1216
25-34	29.9	123/412	33.5	8.63	105/1216
35-44	23.8	98/412	15.5	23.11	281/1216
45-54	18.2	75/412	12.5	29.20	355/1216
55-64	13.6	56/412	5.5	24.01	292/1216
≥65	7.5	31/412	5.5	11.92	145/1216

[a] Only relative data were available.

Figure 4. Distribution of risk categories for invasive pneumococcal disease among *Pneumo Rischio* app and website users (95% CI).



Objective and Subjective Quality of the App as Defined by End Users

A total of 171 participants filled in the MARS user version. The mean age of participants was 34.3 (SD 12.7) years; females were more numerous (107/171, 64.1% women vs 60/171, 35.9% men). A total of 117 (68.4%) respondents were health care professionals. The mean MARS total score was 3.97 (SD 0.45), corresponding to 79.5%.

The mean scores on all MARS subscales exceeded the prespecified threshold of 60%, ranging from 62.8% for the subjective quality subscale to 84.4% for the functionality subscale. With regard to single items, most (15/20) had median scores of 4.0 points, while the median scores were 3.0 points for 4 items and 5.0 points for 1 item (Table 3).

Table 3. Mobile Application Rating Scale mean subscale and median item scores.

Subscale	Subscale scores			Item	Item scores	
	Mean	%	SD		Median	IQR[a]
Engagement	3.75	75.0	0.46	Entertainment	4	3-4
				Interest	4	3-4
				Customization	3	3-3
				Interactivity	3	3-4
				Target group	4	4-5
Functionality	4.22	84.4	0.55	Performance	4	4-5
				Ease of use	5	4-5
				Navigation	4	4-5
				Gestural design	4	4-5
Aesthetics	3.79	75.9	0.55	Layout	4	4-5
				Graphics	4	3-4
				Visual appeal	4	3-4
Information	4.12	82.3	0.51	Quality of information	4	4-4
				Quantity of information	4	4-5
				Visual information	4	4-4
				Credibility of source	4	4-5
Subjective quality	3.14	62.8	0.74	Would recommend the app	4	3-5
				Would use the app in the next 12 months	3	2-3
				Would buy the app	3	1-3
				Overall star rating	4	3-4

[a] IQR: interquartile range.

No significant (t_{169}=0.67, P=.50) between-sex difference emerged in the MARS total scores (mean 4.00, SD 0.40 for men and mean 3.96, SD 0.47 for women). There was a weak positive (r=.16, 95% CI .01-.30) correlation between the score and the age of participants. Non-health care professionals scored significantly (t_{169}=5.59, P<.001) higher than health care professionals (4.20, SD 0.32 vs 3.87, SD 0.46), and the effect size was large (d=0.92, 95% CI 0.58-1.26). Table 4 reports the final multivariable model. The main effect of respondents' professions was a significant (P<.001) predictor of the MARS total score: on average, health care professionals attributed 0.78 (15.6%) fewer points than those outside the health care sector. Moreover, there was a significant (P=.01) interaction between age and profession: with increasing age, health care workers awarded higher scores, while no age-related pattern emerged among non-health care professionals. The model explained 15.5% of variance; its residuals were normally distributed (Shapiro-Wilk test: P=.94) but heteroskedastic (Breusch-Pagan test: P=.02), justifying the use of robust standard errors. The 5 subscale-specific models (Multimedia Appendix 1) yielded very similar results.

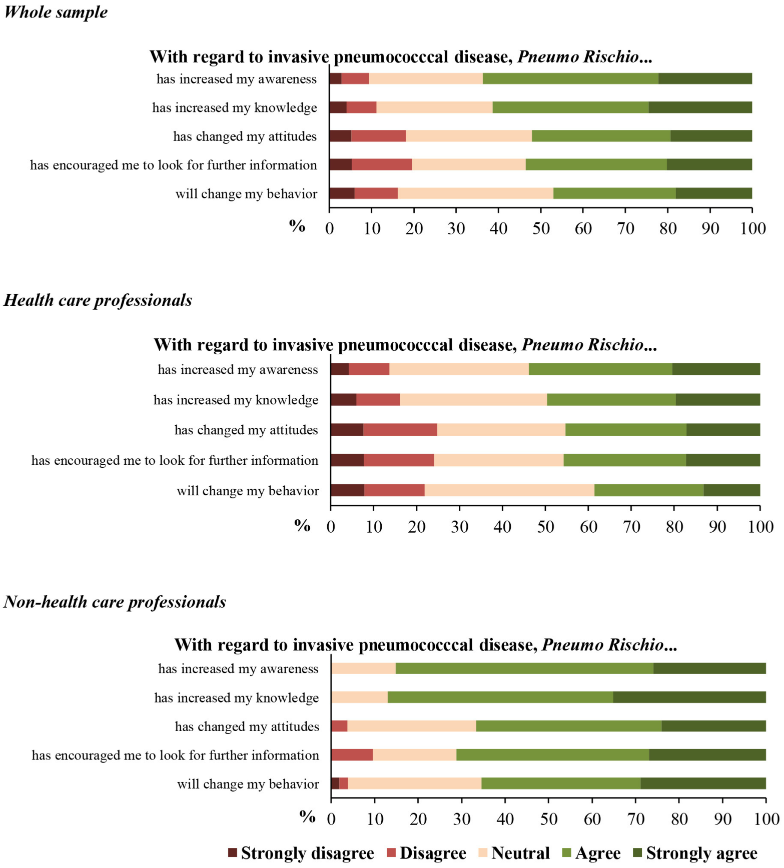
Figure 5 shows response patterns to the app-specific MARS items. More than 60% of respondents attributed a score of at least 4 (ie, "agree" or "strongly agree," which correspond to the light green and dark green areas of the bars in Figure 5) on items regarding app-induced increase in awareness (109/171, 63.7%) and knowledge (105/171, 61.4%), while these proportions were 52.0% (89/171) for items regarding the app's ability to modify attitudes, 53.6% (90/168) for items regarding encouragement of a search for further information, and 47.0% (78/166) for items regarding behavior change. However, like the results of the regression analysis, the perceived usefulness of the app differed by professional category. Comparison of the response categories (ie, strongly agree/agree vs neutral/disagree/strongly disagree) revealed that health care professionals assigned markedly lower scores than did people outside the health care sphere on all items.

Table 4. Final multivariable linear model to predict the Mobile Application Rating Scale total score.

Predictor	*b* (SE)	| t_{165}| (*P* value)
Intercept	4.31 (0.13)	32.43 (<.001)
Sex (female vs male)	–0.01 (0.07)	0.21 (.84)
Age	0.23 (0.24)[a]	0.92 (.36)
Profession (health care vs non-health care)	–0.78 (0.20)	3.97 (<.001)
Age × profession	1.36 (0.50)[a]	2.72 (.01)

[a] Estimates are multiplied by 100.

Figure 5. Response patterns on the app-specific user version of the Mobile Application Rating Scale items.



Discussion

Principal Findings and Comparison With Previous Work

The main novelty of our study is that it took a multidimensional approach to evaluating an eHealth project. The variety of techniques used gave us an opportunity to analyze *Pneumo Rischio* from the point of view of developers, end users, and public health authorities. However, as we believe that in the eHealth era [39] perspectives of different stakeholders are merging, our discussion will not try to distinguish among single actors.

One of our central findings is that marketing activities and continuous monitoring of usage indicators are fundamental in order to increase population exposure to topic-specific eHealth projects such as *Pneumo Rischio*. According to the developers of the Canadian national immunization app [24], the diffusion of an app is a crucial parameter of its success, but this may be challenging. Although the *Pneumo Rischio* app is cost- and registration-free, contains no in-app advertisements, is available in all main Italian app stores, and was developed with the active involvement of experts in pneumococcal diseases (features potentially associated with greater success [40-42]), daily monitoring of the project usage parameters immediately after its launch prompted us to sketch a marketing mix. Indeed, we judged the preintervention number of downloads (n=340) and website visits (n=3994) [14] to be insufficient to reach the main project goal.

We noted that the 1-year life cycle of *Pneumo Rischio* was clearly divided into two stages, a fact undoubtedly attributable to the promotional campaign. Thus, promotional activities were essential to disseminate information on the project. Indeed, without this intervention, the daily number of *Pneumo Rischio* app users would probably not have changed, or may even have decreased, while implementation of the marketing strategy increased app downloads exponentially. Moreover, the creation of a sister website and social media activity may be valuable ways of increasing population exposure.

Our promotional campaign may be regarded as a natural experiment with an interrupted time series design. This quasi-experimental approach is very robust in quantifying the longitudinal effects of an intervention [33]. The results of the segmented regression analysis will allow us both to forecast the effects of future promotional activities and to compare the effects of chronologically distant interventions. Indeed, segmented regression models allow specifications with more than one change point [33].

We observed that the usage parameters of the app and website differed substantially. Indeed, while the *Pneumo Rischio* website had 15.5 times more visitors than the number of app downloads, an average session was much shorter. This finding supports the results obtained by Hearn et al [43], who suggested that an app and a website should be regarded as complementary resources. These authors concluded that an app may help to engage users, while a website is a useful information source [43]. A higher level of engagement (in that downloading an app requires more effort than visiting a website) may partly explain the longer duration of app sessions.

The Photo type of Facebook posts had a significantly higher number of engaged users. This confirms previous findings that these types of posts enhance user interaction and user engagement, and may be seen as a useful metric for social media marketing [30]. By contrast, we did not confirm our hypothesis that Facebook posts containing personal or possessive pronouns or adjectives (eg, “me,” “my”) engaged more users. Although it has been amply demonstrated that personalized messages are more suitable in terms of people’s engagement [44,45], a purely linguistic approach through the use of personal determiners is probably insufficient. A higher level of customization may therefore be more fruitful.

It should, however, be borne in mind that, although social networking helped to promote the app, it may also engender some risks. For instance, about 2 months after its launch, *Pneumo Rischio* was cited and criticized by an antivaccination Facebook group, “Autism and Vaccines,” which as of September 2016 had more than 8600 fans. Specifically, their post questioned “the eulogized scientific method” and urged people “not to fall into the trap of confusing pneumococcus with meningococcus”—presumably a reference to the causative agent of meningitis. This post is typical of the antivaccination movement; skewing science is a frequent claim of antivaccine activists in the era of Web 2.0 [46]. Moreover, the post used the word trap (*tranello*), which has a clear negative connotation. Although this post had no negative impact on the image of *Pneumo Rischio* or its daily downloads (no negative reviews or decrease in the number of users were registered around that week), this type of risk is difficult to detect and manage.

As expected, in terms of age and sex, *Pneumo Rischio* users were not fully representative of Italian adult Web users [47]. Indeed, most app users and Facebook fans were female, while in the Italian context male Internet users are slightly prevalent. On the other hand, the sex distribution of the website users was very close to that of the reference population. With regard to age, we observed different patterns among app and website users and Facebook fans. More than half of app users were adults aged 25-44 years, while 61% of the website visitors were 18-34 years old. Facebook fans were somewhat older, in that three-quarters were 35- to 64-year-olds. Only the distribution of app users over 35 years of age was close to the reference population of adult Web users.

A noteworthy result regards the distribution of IPD risk categories: about two-thirds of users could be defined as being at risk of IPD. While, to the best of our knowledge, there are no Italian data on IPD risk distribution, a large German study [35] found a significantly lower proportion of people at risk of IPD in the general population. In our opinion, several factors may have contributed to the phenomenon observed. First, individual users might have engaged in multiple sessions. For example, they could have given truthful answers (in order to discover their own risk) when filling in the checker the first time, and could subsequently have answered hypothetically (eg, “What would happen if I had…?”). To address this shortcoming, we considered only single events displayed by Google Analytics. A second explanation may lie in the self-diagnoses made by the app users. We tried to prevent this by wording questions in the third person (“Has your doctor ever told you…?”). However, considering the high specificity of the topic, we believe that the most probable cause of the higher risk in our population is linked to the overrepresentation of users who are really at risk of IPD and were able to locate the website or download the app. Indeed, people with any chronic disorder are more likely than healthy individuals to have a health-related app on their mobile phones [48].

The observed patterns of the distributions of app users in terms of sex, age, and IPD risk categories are consistent with the results from an Italian survey on eHealth use [49]. Indeed, Siliquini et al [49] found a higher probability of using the Internet for health-related purposes among females, younger people, and people with chronic conditions. Interestingly, in that study most male eHealth users were young adults aged 18-29 years, while this proportion was highest among females aged 30-41 years. Although males account for a higher proportion of mobile phone owners [50], females are more likely to install a health-related app [48].

While users of the app, website, and Facebook page were not fully representative of Italian Internet users in terms of age and sex, it is largely unknown who downloads vaccination- or disease-specific apps and surfs the Web for immunization-related purposes. To better understand issues concerning the representativeness of our data, it is worth comparing usage patterns of the *Pneumo Rischio* website with those of *VaccinarSì* [51], which is one of the largest immunization-related Web portals in Italy. The results of a 2-year usage study of the *VaccinarSì* project have recently been published [52]. The *Pneumo Rischio* and *VaccinarSì* websites had a similar share of male and female users, a pattern that reflects the national use of the Internet by both sexes, in which males are slightly prevalent. This phenomenon, which still persists even in developed societies, is known as the digital gender divide [53]. Age-class distributions were very similar among users of both websites, while the Facebook fans of *Pneumo Rischio* were older than those of *VaccinarSì*. Another similarity between the two websites regards the mean duration of sessions and the number of pages viewed [52].

It is encouraging that about 5000 *Pneumo Rischio* users declared their intention to request further information on IPD from their physician, and about 2000 sent an email (probably to their GPs) with their detailed IPD risk profile. In the modern digital age, doctor–patient relationships change continuously and become more participatory when a better-informed patient is more closely involved in the decision-making process [54]. In a large representative sample of US physicians, 85% of interviewees had at least one patient who had brought Web-acquired health-related information to a visit in order to ask the doctor's opinion on the matter in question [55]. The role of GPs is therefore crucial in enhancing public awareness of IPD and its prevention, especially among so-called vaccine-hesitant people.

To date (as of September 2016), to the best of our knowledge, our investigation is the largest study to have used the MARS in the community setting. Users rated the app as highly functional; that is, highly performing, easy to use and to navigate, and informative (both qualitatively and quantitatively). These 2 MARS subscales exceeded 80%. Indeed, throughout the process of app development [14], the app's usability properties and its easily comprehensible and evidence-based content were our priority. On the other hand, the subjective quality dimension of the MARS displayed a relatively low mean score, only slightly exceeding 60%. This score was probably lowered by item 19, which concerned the potential purchase of the app; this was the only score with an IQR of 1-3. This may, therefore, suggest that highly targeted and disease-specific apps (especially in the case of diseases with a relatively low incidence) should be free of charge. Indeed, free apps are downloaded much more frequently than paid-for apps, and it has been forecast [56] that, in 2016, 94% of apps will be downloaded free of charge. It is also encouraging that the engagement subscale received a mean score of 75%, since, in addition to the quality of information, interactivity is an important factor that contributes to improving customers' intention in credence goods or services [57].

Health care professionals attributed lower MARS scores to all subscales and app-specific items. The adjusted regression models to predict both MARS total and subscale scores confirmed this finding. It is, therefore, plausible that health care professionals have higher expectations of a health-related app. These lower scores are not surprising, since the app target was the general adult population [14]. However, research has suggested that only half of physicians actually know the term IPD [11]. *Pneumo Rischio* would therefore be useful to many GPs, too.

Study Limitations

In interpreting our results we noted three main limitations. First, we were compelled to use proxy measures of project effectiveness, namely intentions to ask GPs for further information on IPD and number of emails with detailed IPD risk profiles sent. In our opinion, these two indicators approximated project effectiveness better than the number of app downloads, website visits, or Facebook fans. An optimal indicator of *Pneumo Rischio* effectiveness would be the proportion of users who really asked their GP for information on IPD and its prevention, or even were vaccinated against pneumococcus after (and because of) using the project. However, obtaining such data is computationally, economically, and ethically challenging. Second, despite the fact that the professional and user versions of MARS are very similar, only the professional Italian version of MARS has been validated. Considering that the two original English versions of MARS differ only slightly (although the user version does not require training), we supposed that the psychometric properties of the professional Italian version of MARS [36] would be similar to those of the user version. Third, as our sample of participants who filled in the MARS was not representative of the population of Italian adult Internet users, our estimates could differ from the average scores of the reference population. Indeed, given that our convenience sample overrepresented health care professionals, who attributed significantly lower MARS scores than the target population of laypeople, the summary MARS scores could be even higher.

Conclusions

Despite its highly specific topic and somewhat niche nature, *Pneumo Rischio* may be deemed a successful project, as it attracted more than 150,000 users in a 1-year period. Moreover, it was also professionally recognized in an eHealth contest. We therefore hope that our project will contribute to the fight against invasive bacterial diseases.

However, in order to reach their potential end users, such projects should be popularized. Indeed, the process of development of a health-related app should be continuous, and not end with the public release of the app. In our opinion, there is a need to develop a multidimensional framework for assessing health-related apps; this should at least include (1) an evidence base, (2) objective and subjective quality, (3) usefulness, (4) usage indicators, and, if applicable, (5) an outcome assessment.

Acknowledgments

The development of this project was supported by sponsorship from Pfizer S.r.l. The sponsor had no role in the project design, development, or reporting. The authors thank Progetti di Impresa S.r.l. for creating the app, website, and Facebook page and for collecting usage data. The authors thank Dr Bernard Patrick for his linguistic review of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Final multivariable linear models to predict Mobile Application Rating Scale mean subscale scores.

[PDF File (Adobe PDF File), 96KB - jmir_v18i12e316_app1.pdf]

References

1. Pletz MW, Maus U, Krug N, Welte T, Lode H. Pneumococcal vaccines: mechanism of action, impact on epidemiology and adaption of the species. Int J Antimicrob Agents 2008 Sep;32(3):199-206. [doi: 10.1016/j.ijantimicag.2008.01.021] [Medline: 18378430]
2. O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, Hib and Pneumococcal Global Burden of Disease Study Team. Burden of disease caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet 2009 Sep 12;374(9693):893-902. [doi: 10.1016/S0140-6736(09)61204-6] [Medline: 19748398]
3. World Health Organization. Pneumococcal conjugate vaccine for childhood immunization--WHO position paper. Wkly Epidemiol Rec 2007 Mar 23;82(12):93-104 [FREE Full text] [Medline: 17380597]
4. Istituto Superiore di Sanità. Dati di sorveglianza delle malattie batteriche invasive aggiornati al 23 dicembre 2015. 2015. URL: http://www.iss.it/binary/mabi/cont/Report_MBI_20151223_v4.pdf [accessed 2016-06-14] [WebCite Cache ID 6iG8LVM1R]
5. Baxter R, Yee A, Aukes L, Snow V, Fireman B, Atkinson B, et al. Risk of underlying chronic medical conditions for invasive pneumococcal disease in adults. Vaccine 2016 Aug 5;34(36):4293-4297. [doi: 10.1016/j.vaccine.2016.07.003] [Medline: 27396493]
6. Curcio D, Cané A, Isturiz R. Redefining risk categories for pneumococcal disease in adults: critical analysis of the evidence. Int J Infect Dis 2015 Aug;37:30-35 [FREE Full text] [doi: 10.1016/j.ijid.2015.05.003] [Medline: 25997673]
7. Amodio E, Costantino C, Boccalini S, Tramuto F, Maida CM, Vitale F. Estimating the burden of hospitalization for pneumococcal pneumonia in a general population aged 50 years or older and implications for vaccination strategies. Hum Vaccin Immunother 2014;10(5):1337-1342 [FREE Full text] [doi: 10.4161/hv.27947] [Medline: 24577505]
8. Prato R, Tafuri S, Fortunato F, Martinelli D. Why it is still important that countries know the burden of pneumococcal disease. Hum Vaccin 2010 Nov;6(11):918-921. [Medline: 21045538]
9. Istituto Superiore di Sanità. Dati e evidenze disponibili per l'utilizzo dei vaccini anti-pneumococcici nei soggetti a rischio di qualsiasi età e per l?eventuale ampliamento dell?offerta ai soggetti anziani. URL: http://www.epicentro.iss.it/temi/vaccinazioni/pdf/Dati%20e%20evidenze%20vaccini%20antipneumococcici.pdf [accessed 2016-06-14] [WebCite Cache ID 6iG8Sda2D]
10. Williams WW, Lu P, O'Halloran A, Kim DK, Grohskopf LA, Pilishvili T, Centers for Disease Control and Prevention. Surveillance of vaccination coverage among adult populations - United States, 2014. MMWR Surveill Summ 2016 Feb 5;65(1):1-36. [doi: 10.15585/mmwr.ss6501a1] [Medline: 26844596]
11. Lode H, Ludwig E, Kassianos G. Pneumococcal infection--low awareness as a potential barrier to vaccination: results of a European study. Adv Ther 2013 Apr;30(4):387-405. [doi: 10.1007/s12325-013-0025-4] [Medline: 23605248]
12. Nichol KL, MacDonald R, Hauge M. Factors associated with influenza and pneumococcal vaccination behavior among high-risk adults. J Gen Intern Med 1996 Nov;11(11):673-677. [Medline: 9120653]
13. Malerba V, Costantino C, Napoli G, Marchese V, Casuccio A, Tabacchi G, ESCULAPIO Working Group. Antimeningococcal and antipneumococcal vaccination determinants: a European systematic literature review. Epidemiol Prev 2015;39(4 Suppl 1):59-64 [FREE Full text] [Medline: 26499417]
14. Panatto D, Domnich A, Gasparini R, Bonanni P, Icardi G, Amicizia D, et al. Development and preliminary data on the use of a mobile app specifically designed to increase community awareness of invasive pneumococcal disease and its prevention. Hum Vaccin Immunother 2016 Apr 2;12(4):1080-1084. [doi: 10.1080/21645515.2015.1114196] [Medline: 26795065]
15. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
16. Google Play. Pneumo Rischio. Mountain View, CA: Google Inc; 2016. URL: https://play.google.com/store/apps/details?id=com.pneumorischio&hl=it [accessed 2016-11-25] [WebCite Cache ID 6mI20WVua]
17. iTunes. Pneumo Rischio. Cupertino, CA: Apple Inc; 2016. URL: https://itunes.apple.com/it/app/pneumo-rischio/id953486935?mt=8 [accessed 2016-11-25] [WebCite Cache ID 6mI27GSXP]
18. Microsoft Store. Pneumo Rischio. Redmond, WA: Microsoft Corporation; 2016 Jan 14. URL: https://www.microsoft.com/it-it/store/p/pneumo-rischio/9wzdncrcwfph [accessed 2016-11-25] [WebCite Cache ID 6mI2N66Wj]
19. Dipartimento di Scienze della Salute Universita degli Studi di Firenze, Dipartimento di Scienze della Salute Universita degli Studi di Genova. Pneumo Rischio: Calcola il tuo rischio potenziale di contrarre una malattia invasiva da pneumococco.

Modena, Italia: Progetti di Impresa Srl; 2015. URL: http://www.pneumorischio.it/servizi/notizie/notizie_homepage.aspx [accessed 2016-11-17] [WebCite Cache ID 6m5cen7FV]

20. Facebook. Pneumo Rischio. Menlo Park, CA: Facebook; 2016. URL: https://www.facebook.com/Pneumo-Rischio-903906422953296/ [accessed 2016-11-25] [WebCite Cache ID 6mI2avHUI]
21. Plaza I, Demarzo MM, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. JMIR Mhealth Uhealth 2013;1(2):e24 [FREE Full text] [doi: 10.2196/mhealth.2733] [Medline: 25099314]
22. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
23. Yang HC. Bon Appétit for apps: young American consumers' acceptance of mobile applications. J Comput Inform Syst 2015 Dec 10;53(3):85-96. [doi: 10.1080/08874417.2013.11645635]
24. Wilson K, Atkinson KM, Penney G. Development and release of a national immunization app for Canada (ImmunizeCA). Vaccine 2015 Mar 30;33(14):1629-1632. [doi: 10.1016/j.vaccine.2015.02.022] [Medline: 25704801]
25. Panatto D, Gasparini R, Bonanni P, Icardi G, Amicizia D, Domnich A. Disegno e sviluppo di “Pneumo Rischio”: una app come nuovo strumento per valutare il potenziale rischio di contrarre una malattia invasiva da pneumococco. 2015 Presented at: 48th National Congress of SITI; October 14-17, 2015; Milan, Italy.
26. Boccalini S, Bonanni P, Gasparini R, Icardi G, Amicizia D, Domnich A. “Pneumo Rischio”: una nuova app per calcolare il potenziale rischio di contrarre una malattia invasiva da pneumococco. 2015 Presented at: 48th National Congress of SITI; October 14-17, 2015; Milan, Italy.
27. McDonnell DD, Lee H, Kazinets G, Moskowitz JM. Online recruitment of targeted populations: lessons learned from a smoking cessation study among Korean Americans. Soc Marketing Q 2010 Aug 26;16(3):2-22. [doi: 10.1080/15245004.2010.500441]
28. Pletikosa Cvijikj I, Michahelles F. Online engagement factors on Facebook brand pages. Soc Network Analysis Mining 2013 Jan 26;3(4):843-861. [doi: 10.1007/s13278-013-0098-8]
29. 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]
30. Cvijikj I, Michahelles F. A case study of the effects of moderator posts within a Facebook brand page. In: Datta A, Shulman S, Lin SD, Sun A, Lim EP, editors. Social Informatics. Third International Conference, SocInfo 2011. Berlin, Germany: Springer; 2011:161-170.
31. Hamed KH, Ramachandra Rao A. A modified Mann-Kendall trend test for autocorrelated data. J Hydrol 1998 Jan;204(1-4):182-196. [doi: 10.1016/S0022-1694(97)00125-X]
32. Killick R, Eckley IA. changepoint: an R package for changepoint analysis. J Stat Software 2014;58(3):1-19. [doi: 10.18637/jss.v058.i03]
33. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther 2002 Aug;27(4):299-309. [Medline: 12174032]
34. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
35. Theidel U, Kuhlmann A, Braem A. Pneumococcal vaccination rates in adults in Germany: an analysis of statutory health insurance data on more than 850,000 individuals. Dtsch Arztebl Int 2013 Nov 1;110(44):743-750 [FREE Full text] [doi: 10.3238/arztebl.2013.0743] [Medline: 24280430]
36. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. BMC Med Inform Decis Mak 2016;16:83 [FREE Full text] [doi: 10.1186/s12911-016-0323-2] [Medline: 27387434]
37. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
38. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
39. Neuhauser L, Kreps GL. Rethinking communication in the E-health era. J Health Psychol 2003 Jan;8(1):7-23. [doi: 10.1177/1359105303008001426] [Medline: 22113897]
40. Pereira-Azevedo N, Osório L, Cavadas V, Fraga A, Carrasquinho E, Cardoso de Oliveira E, et al. Expert involvement predicts mHealth app downloads: multivariate regression analysis of urology apps. JMIR Mhealth Uhealth 2016;4(3):e86 [FREE Full text] [doi: 10.2196/mhealth.5738] [Medline: 27421338]
41. Ghose A, Han SP. Estimating demand for mobile applications in the new economy. Manage Sci 2014 Jun;60(6):1470-1488. [doi: 10.1287/mnsc.2014.1945]
42. Lee G, Raghu TS. Determinants of mobile apps' success: evidence from the App Store market. J Manage Inf Syst 2014 Nov 10;31(2):133-170. [doi: 10.2753/MIS0742-1222310206]

43. Hearn L, Miller M, Lester L. Reaching perinatal women online: the Healthy You, Healthy Baby website and app. J Obes 2014;2014:573928 [FREE Full text] [doi: 10.1155/2014/573928] [Medline: 24872891]
44. Lee E, Oh SY. To personalize or depersonalize? When and how politicians’ personalized tweets affect the public's reactions. J Commun 2013 Mar 20;62(6):932-949. [doi: 10.1111/j.1460-2466.2012.01681.x]
45. Kalyanaraman S, Sundar SS. The psychological appeal of personalized content in web portals: does customization affect attitudes and behavior? J Commun 2006 Mar;56(1):110-132. [doi: 10.1111/j.1460-2466.2006.00006.x]
46. Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm--an overview of tactics and tropes used online by the anti-vaccination movement. Vaccine 2012 May 28;30(25):3778-3789. [doi: 10.1016/j.vaccine.2011.11.112] [Medline: 22172504]
47. Fondazione UB, Istituto Nazionale di Statistica. Internet@Italia 2014. L'uso di Internet da parte di cittadini e imprese. URL: http://www.istat.it/it/files/2015/12/Internet@Italia2014.pdf [accessed 2016-06-14] [WebCite Cache ID 6iGA8aPsO]
48. Fox S, Duggan M. Mobile health 2012. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2012 Nov 08. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2016-11-17] [WebCite Cache ID 6m5bMHbIe]
49. Siliquini R, Ceruti M, Lovato E, Bert F, Bruno S, De Vito V, et al. Surfing the internet for health information: an italian survey on use and population choices. BMC Med Inform Decis Mak 2011;11:21 [FREE Full text] [doi: 10.1186/1472-6947-11-21] [Medline: 21470435]
50. Poushter J. Smartphone ownership and Internet usage continues to climb in emerging economies. Washington, DC: Pew Research Center; 2016 Feb 22. URL: http://www.pewglobal.org/files/2016/02/pew_research_center_global_technology_report_final_february_22__2016.pdf [accessed 2016-03-15] [WebCite Cache ID 6g2OhVYhl]
51. Vaccinar Si. Rome, Italy: Società Italiana d'Igiene Medicina Preventiva e Sanità Pubblica; 2015. URL: http://www.vaccinarsi.org/ [accessed 2016-11-17] [WebCite Cache ID 6m5dDsKQg]
52. Ferro A, Odone A, Siddu A, Colucci M, Anello P, Longone M, et al. Monitoring the web to support vaccine coverage: results of two years of the portal VaccinarSì. Epidemiol Prev 2015;39(4 Suppl 1):88-93 [FREE Full text] [Medline: 26499422]
53. Bimber B. Measuring the gender gap on the Internet. Soc Sci Q 2000;81(3):868-876.
54. Gerber BS, Eiser AR. The patient physician relationship in the Internet age: future prospects and the research agenda. J Med Internet Res 2001;3(2):E15 [FREE Full text] [doi: 10.2196/jmir.3.2.e15] [Medline: 11720957]
55. Murray E, Lo B, Pollack L, Donelan K, Catania J, Lee K, et al. The impact of health information on the Internet on health care and the physician-patient relationship: national U.S. survey among 1.050 U.S. physicians. J Med Internet Res 2003;5(3):e17 [FREE Full text] [doi: 10.2196/jmir.5.3.e17] [Medline: 14517108]
56. Rivera J, van der Meulen R. Gartner says mobile app stores will see annual downloads reach 102 billion in. Stamford, CT: Gartner, Inc; 2013 Sep 19. URL: http://www.gartner.com/newsroom/id/2592315 [accessed 2016-06-14] [WebCite Cache ID 6iG7ZyupM]
57. Chiu H, Hsieh Y, Kao C. Website quality and customer's behavioural intention: an exploratory study of the role of information asymmetry. Total Quality Manage Business Excellence 2010 Dec 03;16(2):185-197. [doi: 10.1080/14783360500054277]

Abbreviations

GP: general practitioner
IPD: invasive pneumococcal disease
IQR: interquartile range
MARS: Mobile Application Rating Scale

Edited by G Eysenbach; submitted 20.06.16; peer-reviewed by N Bragazzi, Y Albalawi, JR Bautista; comments to author 08.09.16; revised version received 10.10.16; accepted 12.10.16; published 02.12.16.

Please cite as:
Panatto D, Domnich A, Gasparini R, Bonanni P, Icardi G, Amicizia D, Arata L, Carozzo S, Signori A, Bechini A, Boccalini S
An eHealth Project on Invasive Pneumococcal Disease: Comprehensive Evaluation of a Promotional Campaign
J Med Internet Res 2016;18(12):e316
URL: http://www.jmir.org/2016/12/e316/
doi:10.2196/jmir.6205
PMID:27913372

Original Paper

An Ecological Momentary Intervention for Smoking Cessation: Evaluation of Feasibility and Effectiveness

Michael S Businelle[1,2], PhD; Ping Ma[3], PhD; Darla E Kendzor[1,2], PhD; Summer G Frank[2], PhD; Damon J Vidrine[1,2], DrPH; David W Wetter[4], PhD

[1]Department of Family and Preventive Medicine, University of Oklahoma Health Sciences Center, Oklahoma City, OK, United States
[2]Oklahoma Tobacco Research Center, Stephenson Cancer Center, Okahoma City, OK, United States
[3]Division of Population Health, Children's Medical Center, Dallas, TX, United States
[4]Department of Population Health Sciences and the Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Michael S Businelle, PhD
Department of Family and Preventive Medicine
University of Oklahoma Health Sciences Center
655 Research Parkway, Suite 400
Oklahoma City, OK, 73104
United States
Phone: 1 405 271 8001 ext 50460
Fax: 1 405 271 8001
Email: michael-businelle@ouhsc.edu

Abstract

Background: Despite substantial public health progress in reducing the prevalence of smoking in the United States overall, smoking among socioeconomically disadvantaged adults remains high.

Objective: To determine the feasibility and preliminary effectiveness of a novel smartphone-based smoking cessation app designed for socioeconomically disadvantaged smokers.

Methods: Participants were recruited from a safety-net hospital smoking cessation clinic in Dallas, Texas, and were followed for 13 weeks. All participants received standard smoking cessation clinic care (ie, group counseling and cessation pharmacotherapy) and a smartphone with a novel smoking cessation app (ie, Smart-T). The Smart-T app prompted 5 daily ecological momentary assessments (EMAs) for 3 weeks (ie, 1 week before cessation and 2 weeks after cessation). During the precessation period, EMAs were followed by messages that focused on planning and preparing for the quit attempt. During the postcessation period, participant responses to EMAs drove an algorithm that tailored messages to the current level of smoking lapse risk and currently present lapse triggers (eg, urge to smoke, stress). Smart-T offered additional intervention features on demand (eg, one-click access to the tobacco cessation quitline; “Quit Tips” on coping with urges to smoke, mood, and stress).

Results: Participants (N=59) were 52.0 (SD 7.0) years old, 54% (32/59) female, and 53% (31/59) African American, and 70% (40/57) had annual household income less than US $16,000. Participants smoked 20.3 (SD 11.6) cigarettes per day and had been smoking for 31.6 (SD 10.9) years. Twelve weeks after the scheduled quit date, 20% (12/59) of all participants were biochemically confirmed abstinent. Participants responded to 87% of all prompted EMAs and received approximately 102 treatment messages over the 3-week EMA period. Most participants (83%, 49/59) used the on-demand app features. Individuals with greater nicotine dependence and minority race used the Quit Tips feature more than their counterparts. Greater use of the Quit Tips feature was linked to nonabstinence at the 2 (P=.02), 4 (P<.01), and 12 (P=.03) week follow-up visits. Most participants reported that they actually used or implemented the tailored app-generated messages and suggestions (83%, 49/59); the app-generated messages were helpful (97%, 57/59); they would like to use the app in the future if they were to lapse (97%, 57/59); and they would like to refer friends who smoke to use the Smart-T app (85%, 50/59). A minority of participants (15%, 9/59) reported that the number of daily assessments (ie, 5) was “too high.”

Conclusions: This novel just-in-time adaptive intervention delivered an intensive intervention (ie, 102 messages over a 3-week period), was well-liked, and was perceived as helpful and useful by socioeconomically disadvantaged adults who were seeking smoking cessation treatment. Smartphone apps may be used to increase treatment exposure and may ultimately reduce tobacco-related health disparities among socioeconomically disadvantaged adults.

(J Med Internet Res 2016;18(12):e321) doi:10.2196/jmir.6058

KEYWORDS

smartphone; mobile applications; smoking cessation; low income population

Introduction

Tobacco use is the leading cause of preventable death in the United States [1]. Although the prevalence of smoking has declined to 15.2% among US adults who are not living in poverty, 26.3% of those living in poverty are current smokers [2]. Numerous studies have shown that low socioeconomic status (SES) and financial strain are associated with a reduced likelihood of smoking cessation (ie, [3-8]) despite the fact that individuals of low SES are just as likely to make a quit attempt [9,10]. Even when smoking cessation interventions are specifically designed for low SES populations, quit rates have been low (eg, biologically confirmed 7-day point prevalence abstinence rates of 7%-13% at 6-month follow-up [11-13]). As such, innovative smoking cessation interventions are needed to help socioeconomically disadvantaged individuals quit smoking.

Smartphone-based smoking cessation apps could play a significant role in improving cessation rates for current and future generations of smokers. Approximately 779,000 individuals download smoking cessation apps onto personal smartphones each month worldwide [14], and lower SES individuals are the fastest growing group of smartphone owners in the United States [15]. In fact, smartphone ownership more than doubled to 50% between 2011 and 2015 in households earning less than US $30,000 per year [15]. Smartphone apps could offer easily accessible, highly tailored, and intensive interventions at a fraction of the cost of traditional smoking cessation counseling, thereby overcoming many of the barriers that have hampered the use of traditional empirically supported smoking cessation treatments among lower SES individuals [16,17].

Ecological momentary assessment (EMA), in which handheld devices (eg, smartphones) are used to capture moment-to-moment experience, allows for the measurement of phenomena in real time within natural settings [18,19]. EMA data may facilitate a better understanding of the mechanisms involved in successful cessation attempts, those affecting smoking lapses, and those implicated in the transition from lapse to relapse. Although multiple studies have identified momentary predictors of smoking cessation and smoking relapse (eg, [20-24]), to our knowledge, no studies have used a participant’s responses to EMAs to automatically prompt tailored smoking cessation interventions in real time. EMAs are often used to assess individuals at multiple time points throughout a day. Thus, momentary changes in key variables can be tracked and potentially used to initiate interventions as they are needed. Using smartphones to detect high relapse risk situations and automatically deliver tailored smoking cessation interventions may help socioeconomically disadvantaged smokers to quit.

The primary objectives of this study were to evaluate the feasibility and effectiveness of the Smart Treatment app (ie, Smart-T), a novel adjunctive, tailored, smartphone-based smoking cessation intervention for smokers of low SES participating in a smoking cessation program at a safety-net hospital clinic. All study participants received usual tobacco cessation clinic care and a smartphone with the Smart-T app.

Methods

Participants and Procedure

This was a nonrandomized feasibility study in which all participants received the Smart-T intervention. Participants were recruited (June 2014 to May 2015) from an established safety-net hospital smoking cessation clinic in Dallas, Texas. Safety-net hospitals provide health care services regardless of ability to pay and, therefore, primarily serve individuals who are uninsured or receiving Medicaid benefits. Participants were included in the study if they (1) earned a score of ≥4 on the Rapid Estimate of Adult Literacy in Medicine–Short Form (REALM-SF) instrument [25], indicating higher than sixth-grade English literacy level; (2) were willing to quit smoking 7 days after their first clinic visit; (3) were ≥18 years of age; (4) had an expired carbon monoxide (ie, CO) level of ≥8 ppm suggestive of current smoking; (5) were currently smoking ≥5 cigarettes per day; and (6) were willing and able to attend 6 weekly assessment sessions (ie, week −1, quit day, week +1, week +2, week +3, week +4) and the 12-week follow-up session. Participants received US $30 gift cards for completing the week −1, quit day, week +4, and week +12 assessment visits. This study was approved by the institutional review boards at the University of Texas Southwestern Medical Center and the University of Texas School of Public Health.

Individuals attending the orientation visit of the Parkland Smoking Cessation Clinic were provided with detailed information about the study and given the opportunity to have their questions answered in a private room to ensure confidentiality. Written informed consent was obtained. All participants were provided with a smartphone (Samsung, Galaxy Light) for 3 weeks (ie, 1 week before cessation and 2 weeks after cessation). Participants were asked to complete 5 smartphone-prompted EMAs per day (see description of EMA types and items below). All participants were given instructions on how to use the study phone and Smart-T app features at the baseline, quit date, and 1-week follow-up visits. Specifically, participants watched a brief video on a tablet computer, created by the research team, that demonstrated general use of the smartphone and how to access and use Smart-T features. Participants also completed practice EMAs and received hands-on guidance on accessing Quit Tips and other smartphone features (see below). Finally, a link to a brief video tutorial appeared on the home screen of each smartphone so that participants could access smartphone instructions at any time. Participants could use the smartphone to make and receive calls, text, and access the Internet all free of charge. Participants were compensated upon the return of the smartphone based on the percentage of prompted EMAs that were completed over the 3-week EMA period. Specifically, those who completed 50%-74% of assessments received a US $40 gift card, those who completed 75%-89% of assessments received a US $80 gift card, and those who completed 90% or more of their assessments received a US $120 gift card. Participants were not compensated for completing participant-initiated assessments (ie, urge assessments, prequit smoking assessments, lapse assessments) or accessing on-demand features (eg, Quit Tips, Phone a Counselor).

Description of Standard Care

The Parkland Hospital smoking cessation program offers all components of an intensive tobacco treatment recommended by the Clinical Practice Guideline [26], including (1) initial assessment of willingness to participate, (2) the use of multiple types of clinicians (eg, medical, nonmedical), (3) at least four counseling sessions, in an individual or group format, that are greater than 10 minutes in duration, (4) counseling that includes problem-solving, skills training, and social support components, and (5) the opportunity to use effective medications to aid in tobacco cessation (eg, nicotine patch, varenicline). Specifically, smokers were referred (usually by their treatment providers) to the tobacco cessation program. Participants attended one initial orientation and educational session provided by a respiratory therapist, followed by weekly group support sessions facilitated by social workers. Participants were seen individually by a physician (or other prescribing provider) on a regular basis to discuss/prescribe medication and to follow up on participant progress with smoking cessation.

Description of Smart-T App Features

The Smart-T app contains multiple components including an EMA delivery and data transfer system, automated messages, and on-demand content. Each of these features is described below.

Ecological Momentary Assessment

Three types of assessments were used in this study: daily diary, random sampling, and event sampling (ie, precessation smoking, urge, postcessation lapse). Daily diary and random assessments were initiated by the Smart-T app. Specifically, the phone audibly and visually cued each daily diary and random assessment for 60 seconds. If the participant did not respond, the assessment was recorded as missed. Daily diary assessments were completed once every day, 30 minutes after waking. Random assessments were initiated 4 times per day during each participant's normal waking hours. Event sampling assessments were initiated by participants. During the first week of assessment (prequit week), participants were instructed to indicate when they were about to smoke (by clicking the "Record Cigarette" button) immediately before smoking each cigarette. Because the assessment burden would be excessive for heavy smokers if each smoking occasion were assessed, the smartphone randomly sampled up to 2 smoking occasions from each participant per day. After completing the precigarette assessment for selected cigarettes, smokers were instructed to smoke as usual. The phone automatically prompted the postcigarette assessment 15 minutes after the precigarette assessment was completed. During the postquit period, participants were instructed to click the "Urge" button when they had an urge to smoke or they felt like they almost smoked and the "About to Slip" or "Already Slipped" buttons when appropriate during the postquit period. "About to Slip" assessments were followed by a second assessment 15 minutes later to assess whether participants actually smoked. All assessments (excluding daily diary assessments) were expected to require approximately 2-3 minutes to complete. Daily diary assessments were expected to require approximately 5 minutes to complete. All assessments were date and time stamped for future analyses. Assessment items were selected based on their hypothesized relations to smoking behavior and temptation and lapse episodes. The items assessed smoking urges, affect, stress, cigarette availability, recent alcohol use, cessation motivation, and related constructs. This EMA methodology is similar to that developed by Shiffman et al [27,28] and Stone et al [29] and has been used by our research team in many previous studies [21,22,30-34].

Automated Messages

For this study, 4 levels of automated messages were developed, and 1 message was delivered at the end of each EMA. Level 0 messages were pushed during the 1 week prequit period *and* during the postquit period when a participant indicated that he or she lapsed and was no longer interested in quitting smoking. During the prequit week, these messages were not tailored but rather were delivered in a predetermined order. Message topics were primarily motivational in nature and focused on planning and preparing for the quit attempt and benefits of quitting (eg, "It's OK to have mixed feelings about quitting. Don't let that stop you! There will be times that you don't feel like quitting! Stick with it anyway!").

During the 2-week postquit period, participants received individually tailored automated messages based on their EMA responses (see [35] for a complete description of the lapse risk estimator). Level 1 messages were delivered when EMA responses indicated a low level of imminent smoking lapse risk, and message content focused on maintaining abstinence motivation and general cessation advice (eg, seeking social support for cessation, coping with various lapse triggers, and benefits of quitting). Level 2 messages were delivered when EMA responses indicated a high imminent risk of smoking lapse *or* the participant already smoked that day or the day before *or* the participant indicated on the daily diary assessment that he or she had a greater than 25% chance of smoking that day. Level 2 messages were also delivered at the end of participant-initiated Urge EMAs, About to Slip EMAs, and Already Slipped EMAs. Level 2 messages primarily focused on ways to cope with current lapse triggers (ie, reported during the current EMA) and were tailored to the highest rated of 4 current lapse triggers (ie, elevated negative affect/stress, elevated smoking urge, easy access to cigarettes, or low motivation to quit). In instances where multiple triggers were equally highly rated, 1 message was delivered with preference given to negative affect/stress, smoking urge, cigarette availability, and motivation to quit, in the given order. Level 3 messages were delivered after lapse occurred and these motivational messages encouraged a return to abstinence (eg, "A slip is a sign that you need to improve your smoking cessation plan. Think about what went wrong and develop a stronger plan to stay quit. Keep trying and YOU WILL SUCCEED!”).

On-Demand Content

Several Smart-T components were available through the study-provided smartphone 24 hours a day, 7 days a week (see Figure 1). First, the “Phone a Counselor” function/button was programmed to automatically call the free Texas Tobacco Quitline (1-877-Yes-Quit) so that participants could reach a live counselor at any time. Second, the “Quit Tips” function/button accessed a menu of treatment-related messages that focused on general smoking cessation advice, various benefits of quitting, and specific suggestions on how participants might cope with stress, urges, and negative mood (see Figure 2). Third, a “Medications” function/button offered information (eg, common side effects, quit statistics, use instructions) about smoking cessation medications that were regularly prescribed by the Parkland Hospital Smoking Cessation Clinic (see Figure 3).

Figure 1. Smart-T postquit home screen.

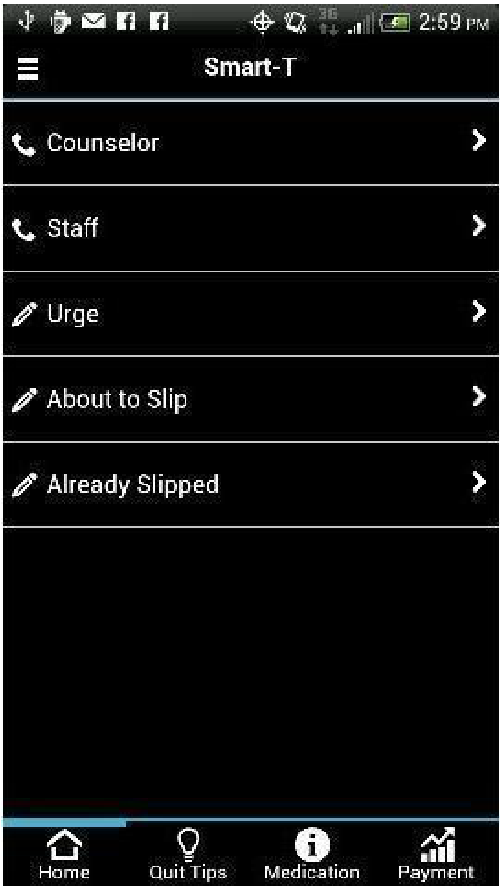


Figure 2. Quit Tips function.

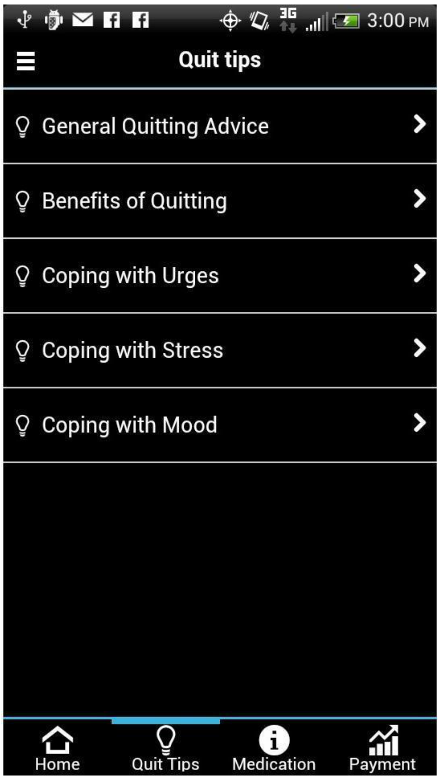
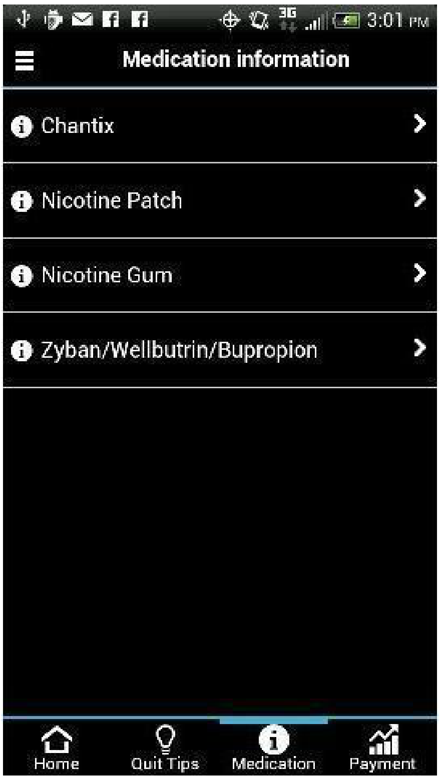


Figure 3. Medication function.



Other App Features

When pressed, the “Call Staff” function/button (see Figure 1) automatically called study staff. Participants were instructed to use this function when they had problems with the phone or the Smart-T app. Pressing the “Payment” button opened a window that indicated the number of EMAs that were prompted and completed and the current level of compensation based on the up-to-the-moment percentage of EMAs completed. In addition, during the prequit period, participants received a daily message that shared the number of days until the participant’s quit date (eg, “You are scheduled to quit smoking in 5 DAYS at 10:00 pm next Sunday night. Developing a plan to quit and taking your medications will GREATLY increase your chances for staying quit.").

Measures

All participants answered demographic questions at the baseline visit including age, sex, race, income, employment status, history of homelessness, and current housing status. In addition, participants completed the Heaviness of Smoking Index (HSI) and the Center for Epidemiological Studies Depression (CES-D) scale at the baseline visit. The HSI is a 2-item measure that is commonly used to assess nicotine dependence [36]. Scores range from 0 to 6 and scores ≥4 indicate moderate to high dependence. The CES-D Short Form is a 10-item measure that is commonly used to assess depressive symptoms [37]. Scores range from 0 to 30 and scores ≥10 indicate clinically significant depression.

At the 2-week follow-up visit, participants were asked a number of questions to gauge their level of satisfaction with and receptiveness to the Smart-T app and app features. Specifically, participants were asked the following: (1) How often they used the automated suggestions that followed each EMA (6-point scale from "Never" to "Always"); (2) Whether the number of EMAs was "Too high," "About right," or "Not enough"; (3) If the EMAs made them more aware of their thoughts, feelings, and behaviors (4-point scale from "Definitely no" to "Definitely yes"); (4) Whether the app helped them to make decisions that were supportive of quitting and staying quit (4-point scale from "Definitely no" to "Definitely yes"); (5) Whether they thought the app was "annoying" (5-point scale from "Not at all" to "Extremely"); (6) If they would recommend the app to a friend (7-point scale ranging from "Extremely unlikely" to "Extremely likely"); and (7) Whether they would be interested in using the app in the future if needed (5-point scale from "Not at all interested" to "Extremely interested"). Participants were also asked if they used each of the on-demand app features. Those who reported using particular features were asked about the usefulness of the feature. Answer options ranged from "Not at all useful" to "Extremely useful" on a 5-point scale. Finally, participants were asked, "At the end of every assessment, the phone automatically offered a tip or suggestion about smoking or smoking cessation. Overall, how helpful were these messages?" Answer options ranged from "Not at all helpful" to "Extremely helpful" on a 5-point scale.

On the quit date, participants were asked if they smoked "even a puff" since 10:00 pm on the night before their quit date visit. Participants were asked if they smoked "even a puff" during the past 7 days at each visit following the scheduled quit date (ie, postquit weeks 1, 2, 3, 4, and 12). Expired breath was tested for CO at each visit using a Vitalograph CO monitor (Lenexa, KS). Self-reported abstinence over the specified time period and a CO reading below 8 ppm (10 ppm on the quit date) were required to be considered abstinent.

Statistical Analyses

Demographic variables, on-demand app feature usage, and participant perceptions of the app are summarized using the mean (SD) and the median for continuous variables and frequency (%) for categorical variables. The frequency with which each participant used each of the app features (ie, Quit Tips, Medication Tips, Phone a Counselor) was calculated, as was the average number of tips seen by each participant during each tip viewing session (ie, viewing intensity). Similar to many count measures, on-demand app use (eg, number of times each tip category was accessed, number of tips seen) was not normally distributed; therefore, negative binomial regression was used to identify differences in app use. These analyses were used to determine if on-demand app use (ie, frequency and intensity of overall Quit Tip and Medication Tip use) was associated with demographic variables and CO confirmed smoking status on the quit date and weeks 2, 4, and 12 postquit visits. When outliers were observed in modeling associations with on-demand feature use, the analysis was repeated after excluding outliers. Data were analyzed using SAS (version 9.4, SAS Institute Inc).

Results

Participants

A total of 61 participants were enrolled in this study. Of these, 2 participants (a 33-year-old African American female and a 54-year-old white male) did not complete any EMAs and were thus excluded from study analyses (sample N=59). On average, participants were 52.0 (SD 7.0) years old, African American (53%, 31/59), and female (54%, 32/59). Most participants were unemployed (78%, 46/59), 70% (40/57) earned less than US $16,000 per year, and half (49%, 29/59) had a history of homelessness. Nearly half (44%, 26/59) of the participants had significant symptoms of depression as measured by the CES-D. Nearly all study participants possessed activated cell phones (93%, 55/59) and 71% (42/59) of all participants possessed activated smartphones. At baseline, participants smoked 20.3 (SD 11.6) cigarettes per day, had been smoking for 31.6 (SD 10.9) years, had low to moderate nicotine dependence (HSI mean 3.5, SD 1.4), and had high levels of expired carbon monoxide (mean 18.6 ppm, SD 13.0).

Attrition and Ecological Momentary Assessment Completion Rates

Most participants attended the quit date (92%, 54/59), week 1 (98%, 58/59), week 2 (98%, 58/59), week 4 (83%, 49/59), and week 12 (78%, 46/59) follow-up visits. Overall, participants were very responsive to prompted EMAs (87% of all prompted EMAs were completed) and EMA completion rates were not related to in-person treatment attendance (*P*=.17). All study phones were returned. Daily diary EMAs were completed in 6.1 (SD 1.9) and 4.8 (SD 1.5) minutes for pre- and postquit assessments, respectively. Random EMAs were completed in 2.3 (SD 0.6) and 2.1 (SD 0.9) minutes for pre- and postquit assessments, respectively. Interestingly, completion rates for the longer daily diary assessments (92% completed) were slightly higher than the shorter random assessments (85% completed). On average, participants self-initiated 15.5 assessments (ie, cigarette, urge, and lapse assessments) and completion times varied by type of participant-initiated EMA. The longest participant-initiated EMA type (ie, Already Slipped assessment) was completed in 4.5 (SD 1.5) minutes on average, while the shortest (ie, Urge assessment) was completed in 2.0

(SD 0.9) minutes. As indicated by *t* tests, there was no relation between cell phone ownership (yes vs no) and EMA completion rates (*P*=.24). However, EMA completion rates differed between smartphone owners and nonowners (*P*=.04), such that smartphone owners completed 10% more assessments than nonowners.

Smart-T Feature Utilization

On-Demand Feature Use

Participants with higher levels of nicotine dependence (ie, higher scores on the HSI) accessed the Quit Tips feature more frequently (6.32 vs 2.13) than those with lower nicotine dependence (*P*<.01). In addition, nonwhite participants accessed the Quit Tips feature twice as often (nonwhite mean 6.53 times vs white mean 3.05 times; *P*=.05) and accessed the Medication Tips nearly 3 times more frequently (nonwhite mean 5.56 vs white mean 1.95; *P*<.01) than white participants. No other demographic variables (eg, income, education) were related to frequency or intensity of Quit Tip or Medication Tip use.

Most participants (83%, 49/59) accessed the Quit Tips feature at least once during the 3-week smartphone use period. Quit Tips were accessed an average of 6.5 (SD 8.0) times per participant (median 4), and an average of 31 (SD 54.1) tips were viewed on each occasion (median 18). When stratified by Quit Tip type, “Coping with Urges” was accessed by the greatest number of participants (n=36), while the highest number of tips viewed per participant per occasion was the General Quitting Advice tip type. The number of times each feature was accessed and number of tips viewed per occasion are reported in Table 1. Means and medians are reported because of outliers who sometimes viewed very large numbers of tips. Most participants (83%, 49/59) accessed the Medication Tips feature on at least one occasion. This feature was accessed 5.2 (SD 5.3) times on average (median 4) and participants viewed 15.4 (SD 31.1) tips per occasion (median 9.8). Interestingly, many participants viewed Medication Tips for medications that were not prescribed for them (see Table 1). Phone log data indicated that very few participants used the “Phone a Counselor” function. Only 13 participants made at least one call to the quitline (ie, 2 minutes or longer) during the 3-week ecological momentary intervention period. A total of 18 calls were made by these 13 participants (mean 8.8 minutes per call; median 3.1 minutes per call).

Table 1. On-demand Quit Tip and Medication Tip use.

Tip type	Number of participants who used the feature	Participants who viewed tips but were not prescribed the medication	Number of times a feature was accessed		Number of tips viewed per occasion	
	n (%) or n	n (%)	Mean (SD)	Median	Mean (SD)	Median
Quit Tips	49 (83)	N/A[a]	6.5 (8.0)	4	31.1 (54.1)	17.8
General Quitting Advice	30	N/A	2.4 (2.6)	2	45.5 (82.4)	15
Benefits of Quitting	21	N/A	1.8 (1.2)	1	26.0 (23.4)	20
Coping with Urges	36	N/A	2.3 (1.7)	2	22.3 (18.9)	16
Coping with Stress	25	N/A	2.4 (2.6)	2	9.9 (7.9)	9
Coping with Mood	31	N/A	2.1 (2.0)	2	13.9 (9.1)	11
Medication Tips	49 (83)		5.2 (5.3)	4	15.4 (31.1)	9.8
Varenicline (Chantix)	33	15 (45)	2.9 (4.1)	2	12.9 (14.2)	10
Nicotine Patch	27	9 (33)	2.9 (2.4)	2	44.4 (164.2)	11
Nicotine Gum	16	12 (75)	1.9 (1.9)	1	4.9 (4.0)	4
Bupropion (Zyban)	26	16 (62)	2.0 (1.5)	1.5	7.0 (8.9)	4

[a] N/A: not applicable.

Automated Messages

On average, participants received 102.1 (SD 23.7) automated intervention messages following EMAs during the 21-day EMA period.

Participant Perceptions of the Smart-T App

Participants answered questions about the usability and helpfulness of the Smart-T app and particular app features 2 weeks after the scheduled quit date (after completing EMAs and using the app for 3 weeks; the phone was returned on this date). Participants reported that the app-generated messages were helpful (97%, 57/59) and encouraged decisions that were supportive of quitting and staying quit (93%, 55/59). Most participants (90%, 53/59) reported that the app made them more aware of their thoughts, feelings, and behaviors, and 83% (49/59) reported that they used the app-generated tailored messages during the postquit period “sometimes, fairly often, very often, or always.” Most participants reported that the number (ie, 5) of daily EMAs that were prompted by the app was “about right” (75%, 44/59) or “not enough” (10%, 6/59), while 15% (9/59) reported that the number of assessments was “too high.” A minority of participants (14%, 8/59) reported that the Smart-T app was “very” or “extremely” annoying. Finally, most participants reported that they would like to use the Smart-T app in the future if needed (97%, 57/59), and 85%

(50/59) would recommend friends who smoke to use the Smart-T app to help them quit. Table 2 displays self-reported use of app features and participant ratings of the usefulness of each feature. Note that only those who reported using each feature were asked about the perceived utility of that feature.

Table 2. Self-reported app feature use and usefulness.

Function	Self-reported use of the function (N=59) n (%)	How helpful or useful was the function?		
		Answer option	n (%)	Mean (SD)
Automated Messages	59 (100)	Not at all helpful	2 (3)	3.8 (1.0)
		Slightly helpful	4 (7)	
		Moderately helpful	15 (25)	
		Very helpful	22 (37)	
		Extremely helpful	16 (27)	
Quit Tips	45 (76)	Not at all useful	0 (0)	3.6 (1.0)
		Slightly useful	9 (20)	
		Moderately useful	10 (22)	
		Very useful	16 (36)	
		Extremely useful	10 (22)	
Medication Tips	38 (64)	Not at all useful	1 (3)	3.5 (1.0)
		Slightly useful	6 (16)	
		Moderately useful	12 (32)	
		Very useful	13 (34)	
		Extremely useful	6 (16)	
Phone a Counselor	10 (17)	Not at all useful	0 (0)	3.1 (1.0)
		Slightly useful	3 (30)	
		Moderately useful	4 (40)	
		Very useful	2 (20)	
		Extremely useful	1 (10)	

Smoking Status and On-Demand Feature Use

A total of 41% (24/59), 17% (10/59), 31% (18/59), 27% (16/59), 22% (13/59), and 20% (12/59) of participants met criteria for point prevalence abstinence at the quit date, week 1, week 2, week 3, week 4, and week 12 follow-up visits, respectively. Participants who did not attend a particular follow-up visit were considered nonabstinent. Notably, the proportion of abstinent participants was significantly higher at the 2-week follow-up compared with the 1-week follow-up (McNemar *P*=<.001).

Negative binomial regression analyses indicated that there was no significant association between the frequency or intensity of tips (ie, Quit Tips and Medication Tips) viewed during the prequit period and biochemically confirmed smoking status on the scheduled quit date (*P* values ≥.40). However, there was a significant relation between the total number of Quit Tips viewed and week 12 smoking status (*P*<.01). Specifically, participants who viewed greater numbers of Quit Tips had a greater likelihood of nonabstinence at the 12-week follow-up visit. Importantly, 2 individuals were identified as extreme outliers (ie, they viewed far greater numbers of Quit Tips compared with other participants). Analyses that excluded these outliers indicated that viewing more Quit Tips was associated with greater likelihood of nonabstinence at the week 2 (*P*=.015), 4 (*P*=.001), and 12 (*P*=.027) postquit visits. No other significant associations were found between frequency or intensity of Medication Tip or Quit Tip use and smoking status at follow-up visits (all *P* values >.09).

Discussion

Principal Findings

Study findings indicate that this first-of-its-kind app offers a feasible way to provide tailored smoking cessation interventions to socioeconomically disadvantaged adults who are seeking to quit smoking. Over a 3-week period, participants received an intensive level of tailored and automated intervention messages (102 messages on average) and most (97%, 57/59) rated these messages as helpful. Furthermore, nearly all participants who accessed the on-demand Quit Tips (100%) and Medication Tips (97%) features rated them as being useful. Most participants (85%, 50/59) reported that they would refer a friend to use the app, and 97% (57/59) reported that they would like to use

Smart-T if they were to lapse in the future. These findings are consistent with smokers’ requests for smoking cessation apps that provide on-demand messages focused on coping with cravings, motivational messages, outcome expectancies, and facts about the effects of smoking [38,39]. The Smart-T app contains each of these features. Importantly, 20% (12/59) of all participants were biochemically confirmed abstinent 12 weeks following the scheduled quit date. This rate of confirmed abstinence is promising as it is higher than what has been commonly reported in other samples of socioeconomically disadvantaged treatment-seeking smokers [11-13,40]. Tailored just-in-time treatments may offer new ways to address disparities in smoking cessation.

Overall, study participants were highly responsive to the 5 automated daily EMAs (87% completed), a rate that is above the EMA compliance benchmark set by Stone and Shiffman [41] and on par with or better than other studies that have collected EMAs in higher SES samples [42-47]. Only 15% (9/59) of all participants reported that the number of prompted EMAs (5 per day) was “too high.” This is a meaningful finding considering that the EMAs required 2 to 6 minutes to complete and could be prompted at any time during the participant’s normal waking hours. Abroms and colleagues [48] reported that 29% of participants indicated that there were “too many” texts in their pilot test of the Text2Quit smoking cessation program. In that study, participants received up to 25 text messages per week (mean 14.5 messages per week) over the first 4 weeks of the program. There are multiple reasons Smart-T participants may have been more accepting of the high number of prompts (ie, 35 per week) compared with the Text2Quit sample. First, Smart-T participants reported that they appreciated the automatic tailored messages that were delivered at the completion of each EMA. This finding is consistent with previous work that has shown that dynamically tailored interventions are more accepted and effective than static messages [49]. Second, Text2Quit participants were asked about their perceptions of the Text2Quit messages after receiving them for 4 weeks, whereas perceptions about the Smart-T app were collected after only 3 weeks. It is possible that favorable perceptions of the Smart-T app may decline with longer periods of use.

Similar to findings from a previous study [50], participants with greater nicotine dependence accessed the Quit Tips feature more times than those who were less dependent. This finding is in contrast to Zeng et al [51] who showed that heavier smoking predicted lower app use. Furthermore, nonwhite participants accessed the Quit Tips and Medication Tips features 2 to 3 times more often than white participants. It is possible that minority smokers and individuals with higher levels of nicotine dependence may have found the app to be more informative and engaging. In fact, many of the Quit Tip messages that were used for this study were originally created for a trial that examined the clinical utility of a culturally tailored palmtop computer–delivered treatment for smoking cessation among African Americans [52].

Most participants (83%, 49/59) used the on-demand Quit Tips and Medication Tips features and these participants viewed an average of 31 (median 18) Quit Tips and 15 (median 10) Medication Tips per viewing session. Participants used these on-demand features in ways that were not anticipated. For instance, it was expected that participants would access tips more frequently and only view a few tips per occasion. There are opposing views about why participants may have accessed such a large number of tips. For example, participants may have chosen to view so many tips per occasion because many tips may not have been relevant to their current situation, and they viewed tips until they found a relevant one. Alternatively, participants may have viewed such a large number of tips because they found the information useful and engaging. High participant ratings for the on-demand tip features may add weight to the latter explanation. Furthermore, it is puzzling why so many participants viewed Medication Tips for medications that were not prescribed to them. Future research should query participants about their rationale for using specific app features.

The frequency, intensity, and types of on-demand Quit Tips that were accessed warrant discussion. More participants used the “Coping with Urges” tips compared with all other types of tips. In addition, participants accessed the “Benefits of Quitting” tips fewer times than the other Quit Tip types. Furthermore, considering each tip viewing occasion, participants chose to view fewer tips that were focused on “Coping with Stress” or “Coping with Mood” compared with other tip types. These findings may provide insights about preferences for specific types of cessation information and could be used to inform future smartphone apps.

It was unexpected that so few participants (22%) would use the “Phone a Counselor” function, which offered a free one-click connection to Texas Tobacco Quitline counselors. While it is unclear why so few participants used this feature, it is possible that access to weekly counseling sessions and on-demand Smart-T Quit Tips supported the belief that accessing the quitline was unnecessary or redundant. Future research should examine if a one-click quitline counseling feature is utilized when paired with stand-alone, automated, message-based smoking cessation apps.

Although the feasibility of the Smart-T app, as an adjunct to in-person treatment, was the primary outcome of interest for this study, exploratory analyses were conducted to examine smoking cessation in this sample. Only 41% (24/59) of the participants quit smoking on their scheduled quit date and only 17% (10/59) maintained abstinence for the entire first postquit week. Interestingly, the 7-day point prevalence abstinence rate nearly doubled to 31% by the second week after cessation. This level of “recycling” or return to abstinence following an initial lapse is unusual [8,53] and may warrant further study. It may be that the Smart-T app helped some participants to overcome early lapse and return to abstinence. Alternatively, this finding may be an artifact of the small sample size of this study. Approximately 20% of this very low SES sample of patients seeking smoking cessation were abstinent at the 12-week follow-up visit. It is notable that biochemically verified abstinence rates were higher than many other interventions that have been conducted with other low SES samples [12,13,40,54,55] and on par with many interventions that have been conducted in more advantaged samples [56,57]. Furthermore, abstinence rates for this study are on par with or better than recently published studies of higher SES smokers seeking cessation that utilized text messaging, which included repeated suggestions to use quitline counseling and nicotine replacement therapy (NRT; [48,58,59]), or a stand-alone smartphone app [60]. This feasibility study is merely a first step, and more vigorous and controlled testing of the utility of the Smart-T app is needed. An ongoing pilot randomized controlled trial will compare the Smart-T app plus NRT to (1) usual in-person smoking cessation treatment (counseling plus NRT) and (2) the National Cancer Institute QuitGuide app plus NRT.

The total number of Quit Tips viewed was related to CO confirmed abstinence status at study follow-up visits. Specifically, individuals who viewed more tips during the 3-week period when participants had access to the Smart-T app were more likely than those who viewed fewer tips to be biochemically confirmed nonabstinent at the week 2, 4, and 12 postcessation follow-up visits. Previous studies that have examined smartphone-based smoking cessation apps have indicated that greater app use is related to increased likelihood for smoking lapse [61], whereas other studies have indicated that greater app use is related to a lower likelihood of lapse [50,60,62]. There are many reasons why participants might have used, or not used, Smart-T app features. First, some abstinent participants may have infrequently used the on-demand app features because of lack of perceived need, while others may have believed that the on-demand app features were integral to their maintenance of abstinence. Second, nonabstinent participants may have lost interest in quitting and thus avoided using the on-demand features, while other nonabstinent participants may have used the on-demand features to prepare them for future quit attempts. Future studies should directly assess the drivers of on-demand feature use so that more effective procedures and features can be created and added to improve cessation apps.

Limitations

The results of this feasibility study should be considered with study limitations. First, this study was not adequately powered to examine the relation between use of Smart-T app features and lapse. Thus, statistical analyses did not control for potential confounders (eg, age, motivation) of the relationship between level of app use and smoking cessation. In addition, all study participants received a smartphone with the Smart-T app; thus, results cannot be used to determine if the Smart-T app improved cessation outcomes beyond the usual safety-net tobacco cessation clinic care. Finally, because of the budget constraints of this feasibility study, participants were only followed up for 12 weeks after their scheduled quit attempt. A well-powered randomized controlled trial that follows up smokers for at least 6 months is needed to adequately examine the utility of the Smart-T app.

Conclusions and Future Directions

In conclusion, this intensive novel smartphone app that tailored intervention messages based on participant-reported situations and symptoms and offered on-demand access to treatment-based messaging was well-used and well-liked in a sample of socioeconomically disadvantaged (eg, 49%, 29/59, reported at least one period of homelessness in their lifetime) smokers seeking cessation treatment at a safety-net hospital. Easily accessible, highly tailored, intensive, well-liked, and low-burden smartphone-based smoking cessation apps may offer new ways to increase treatment exposure and utilization among underserved socioeconomically disadvantaged smokers who have limited access to effective smoking cessation treatments.

Acknowledgments

We thank the staff at the Parkland Hospital Smoking Cessation Clinic for their work and support throughout the data collection portion of this project. In addition, we thank Dr David Balis for his notable efforts that enabled this research. Funding for this research was provided by the American Cancer Society grant MRSGT-12-114-01-CPPB (to MSB). Data analysis and manuscript preparation were additionally supported through MRSGT-10-104-01-CPHPS (to DEK). Additional programming and technological support was provided through the e-Health Technology Program, a resource of the National Institutes of Health through the MD Anderson Cancer Center Support Grant (CA016672) and the Duncan Family Institute for Cancer Prevention and Risk Assessment. The contents of the manuscript are solely the responsibility of the authors and do not necessarily represent the official views of the sponsoring organizations.

Conflicts of Interest

None declared.

References

1. Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. J Am Med Assoc 2004 Mar 10;291(10):1238-1245. [doi: 10.1001/jama.291.10.1238] [Medline: 15010446]
2. Jamal A, Homa DM, O'Connor E, Babb SD, Caraballo RS, Singh T, et al. Current cigarette smoking among adults - United States, 2005-2014. MMWR Morb Mortal Wkly Rep 2015;64(44):1233-1240 [FREE Full text] [doi: 10.15585/mmwr.mm6444a2] [Medline: 26562061]
3. Siahpush M, Carlin JB. Financial stress, smoking cessation and relapse: results from a prospective study of an Australian national sample. Addiction 2006 Jan;101(1):121-127. [doi: 10.1111/j.1360-0443.2005.01292.x] [Medline: 16393198]

4. Fernández E, Schiaffino A, Borrell C, Benach J, Ariza C, Ramon JM, et al. Social class, education, and smoking cessation: long-term follow-up of patients treated at a smoking cessation unit. Nicotine Tob Res 2006 Feb;8(1):29-36. [doi: 10.1080/14622200500264432] [Medline: 16497597]
5. Wetter DW, Cofta-Gunn L, Irvin JE, Fouladi RT, Wright K, Daza P, et al. What accounts for the association of education and smoking cessation? Prev Med 2005 Apr;40(4):452-460. [doi: 10.1016/j.ypmed.2004.07.001] [Medline: 15530598]
6. Fagan P, Shavers VL, Lawrence D, Gibson JT, O'Connell ME. Employment characteristics and socioeconomic factors associated with disparities in smoking abstinence and former smoking among U.S. workers. J Health Care Poor Underserved 2007 Nov;18(4 Suppl):52-72. [doi: 10.1353/hpu.2007.0119] [Medline: 18065852]
7. Businelle MS, Kendzor DE, Reitzel LR, Costello TJ, Cofta-Woerpel L, Li Y, et al. Mechanisms linking socioeconomic status to smoking cessation: a structural equation modeling approach. Health Psychol 2010 May;29(3):262-273 [FREE Full text] [doi: 10.1037/a0019285] [Medline: 20496980]
8. Kendzor DE, Businelle MS, Poonawalla IB, Cuate EL, Kesh A, Rios DM, et al. Financial incentives for abstinence among socioeconomically disadvantaged individuals in smoking cessation treatment. Am J Public Health 2015 Jun;105(6):1198-1205. [doi: 10.2105/AJPH.2014.302102] [Medline: 25393172]
9. Kotz D, West R. Explaining the social gradient in smoking cessation: it's not in the trying, but in the succeeding. Tob Control 2009 Feb;18(1):43-46 [FREE Full text] [doi: 10.1136/tc.2008.025981] [Medline: 18936053]
10. Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and smoking: a review. Ann N Y Acad Sci 2012 Feb;1248:107-123. [doi: 10.1111/j.1749-6632.2011.06202.x] [Medline: 22092035]
11. Froelicher ES, Doolan D, Yerger VB, McGruder CO, Malone RE. Combining community participatory research with a randomized clinical trial: the protecting the hood against tobacco (PHAT) smoking cessation study. Heart Lung 2010;39(1):50-63. [doi: 10.1016/j.hrtlng.2009.06.004] [Medline: 20109986]
12. Hahn EJ, Rayens MK, Chirila C, Riker CA, Paul TP, Warnick TA. Effectiveness of a quit and win contest with a low-income population. Prev Med 2004 Sep;39(3):543-550. [doi: 10.1016/j.ypmed.2004.02.012] [Medline: 15313094]
13. Okuyemi KS, James AS, Mayo MS, Nollen N, Catley D, Choi WS, et al. Pathways to health: a cluster randomized trial of nicotine gum and motivational interviewing for smoking cessation in low-income housing. Health Educ Behav 2007 Feb;34(1):43-54. [doi: 10.1177/1090198106288046] [Medline: 16778147]
14. Abroms LC, Lee WJ, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
15. Pew Research Center. US Smartphone Use in 2015. 2015 May 01. URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2016-12-06] [WebCite Cache ID 6mYwi7RuY]
16. Siahpush M, Yong H, Borland R, Reid JL, Hammond D. Smokers with financial stress are more likely to want to quit but less likely to try or succeed: findings from the International Tobacco Control (ITC) Four Country Survey. Addiction 2009 Aug;104(8):1382-1390 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02599.x] [Medline: 19438837]
17. Copeland AL, Businelle MS, Stewart DW, Patterson SM, Rash CJ, Carney CE. Identifying barriers to entering smoking cessation treatment among socioeconomically disadvantaged smokers. J Smok Cessat 2010;5:164-171. [doi: 10.1375/jsc.5.2.164]
18. Shiffman S, Hufford M, Hickcox M, Paty JA, Gnys M, Kassel JD. Remember that? A comparison of real-time versus retrospective recall of smoking lapses. J Consult Clin Psychol 1997 Apr;65(2):292-300. [Medline: 9086693]
19. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient non-compliance with paper diaries. Br Med J 2002 May 18;324(7347):1193-1194 [FREE Full text] [Medline: 12016186]
20. Shiffman S, Balabanis MH, Paty JA, Engberg J, Gwaltney CJ, Liu KS, et al. Dynamic effects of self-efficacy on smoking lapse and relapse. Health Psychol 2000 Jul;19(4):315-323. [Medline: 10907649]
21. Reitzel LR, Kendzor DE, Nguyen N, Regan SD, Okuyemi KS, Castro Y, et al. Shelter proximity and affect among homeless smokers making a quit attempt. Am J Health Behav 2014 Mar;38(2):161-169 [FREE Full text] [doi: 10.5993/AJHB.38.2.1] [Medline: 24629545]
22. Watkins KL, Regan SD, Nguyen N, Businelle MS, Kendzor DE, Lam C, et al. Advancing cessation research by integrating EMA and geospatial methodologies: associations between tobacco retail outlets and real-time smoking urges during a quit attempt. Nicotine Tob Res 2014 May;16(Suppl 2):S93-101 [FREE Full text] [doi: 10.1093/ntr/ntt135] [Medline: 24057995]
23. Piasecki TM, Jorenby DE, Smith SS, Fiore MC, Baker TB. Smoking withdrawal dynamics: II. Improved tests of withdrawal-relapse relations. J Abnorm Psychol 2003 Feb;112(1):14-27. [Medline: 12653410]
24. Businelle MS, Ma P, Kendzor DE, Reitzel LR, Chen M, Lam CY, et al. Predicting quit attempts among homeless smokers seeking cessation treatment: an ecological momentary assessment study. Nicotine Tob Res 2014 Oct;16(10):1371-1378 [FREE Full text] [doi: 10.1093/ntr/ntu088] [Medline: 24893602]
25. Arozullah AM, Yarnold PR, Bennett CL, Soltysik RC, Wolf MS, Ferreira RM, et al. Development and validation of a short-form, rapid estimate of adult literacy in medicine. Med Care 2007 Nov;45(11):1026-1033. [doi: 10.1097/MLR.0b013e3180616c1b] [Medline: 18049342]

26. Fiore M, Jaen C, Baker T, Bailey W, Benowitz N, Curry S. Treating Tobacco Use and Dependence: 2008 Update. In: Clinical Practice Guideline. Rockville, MD: US Department of Health and Human Services. Public Health Service; May 2008.
27. Shiffman S, Paty JA, Gnys M, Kassel JA, Hickcox M. First lapses to smoking: within-subjects analysis of real-time reports. J Consult Clin Psychol 1996 Apr;64(2):366-379. [Medline: 8871421]
28. Shiffman S, Hickcox M, Paty JA, Gnys M, Richards T, Kassel JD. Individual differences in the context of smoking lapse episodes. Addict Behav 1997;22(6):797-811. [Medline: 9426798]
29. Stone AA, Schwartz JE, Neale JM, Shiffman S, Marco CA, Hickcox M, et al. A comparison of coping assessed by ecological momentary assessment and retrospective recall. J Pers Soc Psychol 1998 Jun;74(6):1670-1680. [Medline: 9654765]
30. Businelle MS, Kendzor DE, Cuate EL, Kesh A, Poonawalla IB. Using smart phones to collect momentary data among sheltered homeless smokers. 2013 Presented at: Annual Conference for the Society for Research on Nicotine and Tobacco; 2013; Boston, Massachusetts.
31. Businelle MS, Lam CY, Kendzor DE, Cofta-Woerpel L, McClure JB, Cinciripini PM, et al. Alcohol consumption and urges to smoke among women during a smoking cessation attempt. Exp Clin Psychopharmacol 2013 Feb;21(1):29-37 [FREE Full text] [doi: 10.1037/a0031009] [Medline: 23379613]
32. Lam CY, Businelle MS, Aigner CJ, McClure JB, Cofta-Woerpel L, Cinciripini PM, et al. Individual and combined effects of multiple high-risk triggers on postcessation smoking urge and lapse. Nicotine Tob Res 2014 May;16(5):569-575 [FREE Full text] [doi: 10.1093/ntr/ntt190] [Medline: 24323569]
33. Lam CY, Businelle MS, Cofta-Woerpel L, McClure JB, Cinciripini PM, Wetter DW. Positive smoking outcome expectancies mediate the relation between alcohol consumption and smoking urge among women during a quit attempt. Psychol Addict Behav 2014 Mar;28(1):163-172. [doi: 10.1037/a0034816] [Medline: 24731113]
34. Kendzor DE, Shuval K, Gabriel KP, Businelle MS, Ma P, High RR, et al. Impact of a mobile phone intervention to reduce sedentary behavior in a community sample of adults: A quasi-experimental evaluation. J Med Internet Res 2016;18(1):e19 [FREE Full text] [doi: 10.2196/jmir.5137] [Medline: 26810027]
35. Businelle MS, Ma P, Kendzor DE, Frank SG, Wetter DW, Vidrine DJ. Using intensive longitudinal data collected via mobile phone to detect imminent lapse in smokers undergoing a scheduled quit attempt. J Med Internet Res 2016 Oct 17;18(10):e275. [doi: 10.2196/jmir.6307] [Medline: 27751985]
36. Kozlowski LT, Porter CQ, Orleans CT, Pope MA, Heatherton T. Predicting smoking cessation with self-reported measures of nicotine dependence: FTQ, FTND, and HSI. Drug Alcohol Depend 1994 Feb;34(3):211-216. [Medline: 8033758]
37. Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression Scale (CES-D). Arch Intern Med 1999;159(15):1701-1704. [Medline: 10448771]
38. Bock BC, Heron KE, Jennings EG, Magee JC, Morrow KM. User preferences for a text message-based smoking cessation intervention. Health Educ Behav 2013 Apr;40(2):152-159. [doi: 10.1177/1090198112463020] [Medline: 23086555]
39. Naughton F, Prevost AT, Gilbert H, Sutton S. Randomized controlled trial evaluation of a tailored leaflet and SMS text message self-help intervention for pregnant smokers (MiQuit). Nicotine Tob Res 2012 May;14(5):569-577. [doi: 10.1093/ntr/ntr254] [Medline: 22311960]
40. Okuyemi KS, Goldade K, Whembolua G, Thomas JL, Eischen S, Sewali B, et al. Motivational interviewing to enhance nicotine patch treatment for smoking cessation among homeless smokers: a randomized controlled trial. Addiction 2013 Jun;108(6):1136-1144 [FREE Full text] [doi: 10.1111/add.12140] [Medline: 23510102]
41. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [Medline: 12173681]
42. Yeh VM, McCarthy DE, Baker TB. An ecological momentary assessment analysis of prequit markers for smoking-cessation failure. Exp Clin Psychopharmacol 2012 Dec;20(6):479-488 [FREE Full text] [doi: 10.1037/a0029725] [Medline: 22924702]
43. Schwartz JE, Neale J, Marco C, Shiffman SS, Stone AA. Does trait coping exist? A momentary assessment approach to the evaluation of traits. J Pers Soc Psychol 1999 Aug;77(2):360-369. [Medline: 10474211]
44. Epstein DH, Willner-Reid J, Vahabzadeh M, Mezghanni M, Lin J, Preston KL. Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use. Arch Gen Psychiatry 2009 Jan;66(1):88-94 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2008.509] [Medline: 19124692]
45. Berkman ET, Dickenson J, Falk EB, Lieberman MD. Using SMS text messaging to assess moderators of smoking reduction: Validating a new tool for ecological measurement of health behaviors. Health Psychol 2011 Mar;30(2):186-194 [FREE Full text] [doi: 10.1037/a0022201] [Medline: 21401252]
46. Köblitz AR, Magnan RE, McCaul KD, O'Neill HK, Crosby R, Dillard AJ. Smokers' thoughts and worries: a study using ecological momentary assessment. Health Psychol 2009 Jul;28(4):484-492. [doi: 10.1037/a0014779] [Medline: 19594273]
47. Shiffman S, Paty J. Smoking patterns and dependence: contrasting chippers and heavy smokers. J Abnorm Psychol 2006 Aug;115(3):509-523. [doi: 10.1037/0021-843X.115.3.509] [Medline: 16866591]
48. Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012;17(Suppl 1):44-53 [FREE Full text] [doi: 10.1080/10810730.2011.649159] [Medline: 22548598]

49. 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]
50. Wetter DW, McClure JB, Cofta-Woerpel L, Costello TJ, Reitzel LR, Businelle MS, et al. A randomized clinical trial of a palmtop computer-delivered treatment for smoking relapse prevention among women. Psychol Addict Behav 2011 Jun;25(2):365-371 [FREE Full text] [doi: 10.1037/a0022797] [Medline: 21500879]
51. Zeng EY, Vilardaga R, Heffner JL, Mull KE, Bricker JB. Predictors of utilization of a novel smoking cessation smartphone app. Telemed J E Health 2015 Dec;21(12):998-1004 [FREE Full text] [doi: 10.1089/tmj.2014.0232] [Medline: 26171733]
52. Businelle MS, Kendzor DE, Costello TJ, Cofta-Woerpel L, Li Y, Mazas CA, et al. Light versus heavy smoking among African American men and women. Addict Behav 2009 Feb;34(2):197-203 [FREE Full text] [doi: 10.1016/j.addbeh.2008.10.009] [Medline: 18976867]
53. Gilpin EA, Pierce JP, Farkas AJ. Duration of smoking abstinence and success in quitting. J Natl Cancer Inst 1997 Apr 16;89(8):572-576 [FREE Full text] [Medline: 9106646]
54. Businelle MS, Kendzor DE, Kesh A, Cuate EL, Poonawalla IB, Reitzel LR, et al. Small financial incentives increase smoking cessation in homeless smokers: a pilot study. Addict Behav 2014 Mar;39(3):717-720. [doi: 10.1016/j.addbeh.2013.11.017] [Medline: 24321696]
55. Christiansen BA, Reeder KM, TerBeek EG, Fiore MC, Baker TB. Motivating low socioeconomic status smokers to accept evidence-based smoking cessation treatment: a brief intervention for the community agency setting. Nicotine Tob Res 2015 Aug;17(8):1002-1011 [FREE Full text] [doi: 10.1093/ntr/ntu345] [Medline: 26180226]
56. Myung S, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
57. 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]
58. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
59. Bennett DA, Emberson JR. Text messaging in smoking cessation: the txt2stop trial. Lancet 2011 Jul 2;378(9785):6-7. [doi: 10.1016/S0140-6736(11)60882-9] [Medline: 21722951]
60. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 1;143:87-94. [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
61. Balmford J, Borland R. How do smokers use a smoking cessation text messaging intervention? Nicotine Tob Res 2014 Dec;16(12):1586-1592. [doi: 10.1093/ntr/ntu111] [Medline: 25031312]
62. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE. Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed J E Health 2014 Mar;20(3):206-214 [FREE Full text] [doi: 10.1089/tmj.2013.0169] [Medline: 24350804]

Abbreviations

CES-D: Center for Epidemiological Studies Depression
EMA: ecological momentary assessment
HSI: Heaviness of Smoking Index
NRT: nicotine replacement therapy
SES: socioeconomic status

Edited by G Eysenbach; submitted 31.05.16; peer-reviewed by JC Chen, V Mevsim, J Walker, P Keller; comments to author 25.09.16; revised version received 19.10.16; accepted 30.11.16; published 12.12.16.

Please cite as:
Businelle MS, Ma P, Kendzor DE, Frank SG, Vidrine DJ, Wetter DW
An Ecological Momentary Intervention for Smoking Cessation: Evaluation of Feasibility and Effectiveness
J Med Internet Res 2016;18(12):e321
URL: http://www.jmir.org/2016/12/e321/
doi:10.2196/jmir.6058
PMID:27956375

Original Paper

A Diabetes Self-Management Program: 12-Month Outcome Sustainability From a Nonreinforced Pragmatic Trial

Kate Lorig[1], DrPH; Philip L Ritter[1], PhD; Ralph M Turner[2], PhD; Kathleen English[3], MBA; Diana D Laurent[1], MPH; Jay Greenberg[4], ScD

[1]Stanford Patient Education Research Center, Medicine, Stanford University, Palo Alto, CA, United States
[2]HealthCore, Wilmington, DE, United States
[3]Anthem Inc, Indianapolis, IN, United States
[4]NCOA Services, LLC, Arlington, VA, United States

Corresponding Author:
Philip L Ritter, PhD
Stanford Patient Education Research Center
Medicine
Stanford University
1000 Welch Rd
Palo Alto, CA, 94306
United States
Phone: 1 650 725 2873
Fax: 1 650 725 9422
Email: philr@stanford.edu

Abstract

Background: Diabetes self-management education has been shown to be effective in controlled trials. The 6-week Better Choices, Better Health-Diabetes (BCBH-D) self-management program was also associated with an improvement in health outcomes in a 6-month translation study.

Objective: The objective of this study was to determine whether a national translation of the BCBH-D self-management program, offered both Web-based and face-to-face, was associated with improvements in health outcomes (including HbA1c) and health behaviors (including recommended medical tests) 1 year after intervention

Methods: Web-based programs were administered nationally, whereas face-to-face workshops took place in Atlanta, Indianapolis, and St Louis. Self-report questionnaires were either Web-based or administered by mail, at baseline and 1 year, and collected health and health-behavior measures. HbA1c blood samples were collected via mailed kits. A previous 6-month study found statistically significant improvements in 13 of 14 outcome measures, including HbA1c. For this study, paired t test compared baseline with 1-year outcomes. Subgroup analyses determined whether participants with specific conditions improved (high HbA1c, depression, hypoglycemia, nonadherence to medication, no aerobic exercise). The percentage of participants with improvements in effect size of at least 0.4 in at least 1 of the 5 measures was calculated.

Results: A total of 857 participants with 1-year data (69.7% of baseline participants) demonstrated statistically significant 1-year improvements in 13 of 15 outcome measures; 79.9% (685/857) of participants showed improvements in effect size of 0.4 or greater in at least 1 of the 5 criterial measures.

Conclusions: Participants had small but significant benefits in multiple measures. Improvements previously noted at 6 months were maintained or amplified at 1 year.

(J Med Internet Res 2016;18(12):e322) doi:10.2196/jmir.6484

KEYWORDS

patient education; self-management; type 2 diabetes

Introduction

Background

Although Healthy People 2020 recommends diabetes education, less than 7% of people with diabetes report receiving formal diabetes education in the year following diagnosis [1]. Diabetes education may also help with achieving many of the Healthcare Effectiveness Data and Information Set (HEDIS) measures [2]. Diabetes self-management education aims to increase healthful behaviors while reducing HbA1c. In a recently published paper, we reported on the 6-month outcomes for a diabetes self-management intervention offered both face-to-face and Web-based [3]. Study participants improved healthful behaviors, medication adherence, hypoglycemia, and HbA1c. More than 70% of study participants made a positive improvement (effect size of 0.4 or greater) in 1 or more of 5 outcome variables (HbA1c, frequency of hypoglycemia, medication adherence, completing recommended screenings, and exercise). In this paper, we report on the 1-year outcomes for the same study.

Hypotheses

The study aimed at testing the following hypotheses:

1. Improvements in HbA1c, health indicators (hypoglycemic symptoms and depression), and health behaviors (exercise, medication adherence, receiving recommended tests) between baseline (preintervention) and 1 year.

2. Changes would meet or exceed Healthy People diabetes recommendation for percentage of population with HbA1c above 9% and below 7% and percentage of population receiving microalbumin, foot, and eye examinations.

3. Effectiveness would be independent of the mode of delivery.

4. Both older (65 plus) and younger participants would benefit.

5. Participants with baseline HbA1c ≥9.0, Patient Health Questioinaire (PHQ-8) depression ≥10.0, 2 or more hypoglycemia symptoms, medication nonadherence, or no aerobic exercise would have clinically significant improvements in these variables.

6. Moderate effect size (0.4) improvements would be found for the majority of participants in reducing 1 of more of the aforementioned variables.

Methods

Intervention

Better Choices, Better Health-Diabetes (BCBH-D) was developed for people with type 2 diabetes. Both versions of the program (face-to-face and Web-based) meet the American Association of Diabetes Educators (AADE) standards for diabetes self-management and support [4], have been shown to be effective in previous randomized trials [5,6], have the same content, and are designed to enhance self-efficacy [7]. Content includes healthy eating, exercise, understanding glucose monitoring, communicating with family, friends, and the health care system, hypoglycemia, depression, difficult emotions, sick days, medication management, problem solving, decision making, and action planning. Both are 6 weeks, have 2 peer facilitators, and have standardized facilitator training. Neither program had any reinforcement beyond the initial intervention. Both have been described in detail elsewhere [3]. The study was approved by the Stanford and New England institutional review boards and has 5 collaborators, Anthem Inc, the National Council on Aging, Stanford School of Medicine, the National Council of Young Men's Christian Association of the United States of America, and OASIS Health.

Recruitment

Because this was a pragmatic trial designed to be offered in real-world settings, there were few inclusion (have type 2 diabetes and be covered by an Anthem-affiliated health plan) or exclusion (currently pregnant or in chemotherapy or radiation treatment for cancer) criteria. Unlike most efficacy trials, symptom severity was not an inclusion criterion.

Web-based participants were recruited in 2013 and 2014 by email from their employers or emails or phone calls from an Anthem plan. Both commercial and Medicare-Advantage participants were eligible. Potential study participants went to the recruitment website, completed screening, and completed an informed consent and baseline questionnaire.

Face-to-face participants were recruited through mailings, flyers in workplaces or physicians' offices, case managers, and automated telephone calls. Face-to-face programs were available in Atlanta, Indianapolis, and St. Louis. A small percentage of the community participants were not covered by Anthem plans. All other screening criteria were the same for Web-based and face-to-face participants.

Data Collection

Data were collected using self-report, validated questionnaires at baseline and 12 months. We asked participants to furnish a sample of blood, although this was not required for program or study participation; nor were participants disqualified if they failed to return their samples. Consenting participants were sent HbA1c test kits. These were returned to investigators, bar-coded to avoid disclosing PHI, and then sent to CoreMedica, a Clinical Laboratory Improvement Amendments (CLIA) certified lab [8]. Participants and their physicians were sent results. Because CoreMedica recalibrated its measurements in June 2014 increasing all values by roughly 0.4, all measures prior to June were adjusted upward by 0.4.

Measures

Measures were chosen to be of interest to patients, providers, and the health care system. Demographic variables included age, gender, race, ethnicity, education, and marital status. Participant also reported other diseases. Previously validated outcome measures are described in detail elsewhere [3,9,10].

Health indicators included self-rated health, which is a single item from the National Health Examination Survey [11], the PHQ-8 depression scale [12], the Illness Intrusiveness Scale, which measures role function [13], and the hypoglycemic symptoms scale [14]. Fatigue and sleep are each single-item visual numeric scales [15]. Behavior indicators included communication with physicians, a scale asking how often patients ask questions and discuss problems with their health care provider [10], minutes of aerobic exercise per week [10], and the Morisky Medication Adherence scale [16]. We also asked participants if they had eye, foot, cholesterol, and kidney examinations in the last year or, on follow-up questionnaires, in the last 6 months.

Data Analysis

Pragmatic studies, because of study-population heterogeneity, present unique methodological challenges. In many diabetes studies, subjects are chosen, for example, because of high HbA1c or depression. In this study, no such screening occurred, resulting in greater heterogeneity for the key outcome variables. Not all participants have the same problems, and some have no problems. Consequently, we conducted several types of analyses. The first was a descriptive analysis of the participants and their engagement with the program. The second, or classic, set of analyses determined the 12-month changes and significance for the population as a whole. The third, or subset, analyses examined only those who demonstrated problems in specific variables of interest, for example, high HbA1c or low adherence to taking medications. Having a problem was decided either by a standard criteria such as 10 or above on a PHQ-8 being highly suggestive of depression [12] or by specific scores on the adherence and hypoglycemia scales suggested by the authors and previous publications. For exercise, we chose those not exercising at baseline, and for HbA1c, we chose 9, as this is the level discussed in the Healthy People 2020 goals [17]. A fourth analysis sought to reconcile the second and third analyses by examining the percentage of the total population who achieved a moderate benefit (0.4 effect size) in at least 1 of the 5 variables of interest.

To help determine the likelihood of bias caused by attrition, we compared baseline scores of participants who failed to complete with those who completed 12-month questionnaires.

Univariate statistics describe demographic and engagement characteristics. Independent-sample *t* tests compare demographic and baseline outcome variables between those who failed to complete 12-month follow-up questionnaires and those who completed them. Paired *t* tests examined changes between baseline and 12 months and if these differed significantly from a null hypothesis of zero change (hypothesis 1).

For those who had had no examinations (eye, foot, cholesterol, or kidney) in the year prior to entry, we calculated the percentage that had examinations in the 12 months following baseline.

To compare effectiveness by mode of delivery, we used independent- sample *t* tests to compare change scores between Web-based and face-to-face participants (hypothesis 3). Similarly, we compared 1-year changes for older (65 plus) and younger participants (hypothesis 4).

Subgroup analyses were conducted for participants with specific conditions as described above: HbA1c above 9; clinical depression (PHQ-8 of 10 or above [12]); at least two symptoms of hypoglycemia; low medication adherence; and no exercise at baseline. For each measure, we report the mean change of the group and the percentage that no longer had the negative indication (hypothesis 5). In addition, to examine the possibility that results were due to regressions to the mean, we calculated the change scores for the subsets of the sample that did not meet each of the 5 negative criteria.

To determine the proportion of all participants who benefited on at least 1 of the 5 indicators, we calculated the percentage who improved by an effect size of at least 0.4 (hypothesis 6). We examined the relationship between the number of criterial indicators and the number of improvements using Pearson correlations.

Results

Participants

A total of 4639 potential participants (509 face-to-face and 4130 Web-based) left contact information. All were invited to fill out a screener and, if eligible, complete consent and the baseline questionnaire. People only became study participants when they attended the first workshop or logged on to the first session. Of these, 1229 (229 face-to-face and 1000 Web-based) completed the process, were eligible, and became study participants. The majority of those not completing the process did not have Anthem-affiliated plans or failed to complete the screener, consent, or questionnaire (Figures 1 and 2). Of those starting the program, 687 and 170 (69.7%) completed 12-month questionnaires. About 80% (957/1229) of the participants successfully completed a baseline HbA1c. Of these, 55.0% (526/957) completed 12-month HbA1cs.

Table 1 shows demographic characteristics and baseline outcome scores. Relative to the face-to-face cohort, the Web-based participants were more likely to be male, more likely to be married, and less likely to be minority. They also were younger. All differences were statistically significant. There was 1 statistically significant difference between the 2 groups among baseline outcome measures. The face-to-face participants had better initial communication with their physicians (*P*=.03).

Table 1. Baseline characteristics of participants with 12-month data.

Variable	Face-to-face (n=170)	Web-based (n=687)	All (N=857)
Male, n (%)	47 (27.6)	238 (34.6)	285 (33.3)
Education in years, mean (SD, range)	15.2 (2.91, 8-23)	15.6 (2.78, 10-23)	15.5 (2.81, 8-23)
Married, %	89 (52.3)	518 (75.5)	607 (70.9)
Non-Hispanic white, %	114 (67.1)	524 (83.5)	638 (74.4)
Black, %	49 (28.8)	57 (8.3)	106 (12.4)
Hispanic, %	3 (1.8)	60 (8.7)	63 (7.4)
Age in years, mean (SD, range)	65.6 (9.95, 28-95)	55.8 (8.62, 26-91)	57.7 (9.71, 26-95)
Number of other chronic conditions, mean (SD, range)	1.65 (1.37, 0-8)	1.40 (1.15, 0-6)	1.45 (1.19, 0-8)
Medicare, %	97 (57.6)	48 (7.0)	145 (16.9)
Private insurance, %	124 (72.9)	687 (100)	811 (94.5)
HbA1c ↓[a], n=732, mean (SD, range)	7.79 (1.55, 5-15.2)	8.02 (1.28, 5-12.1)	8.04 (1.44, 5-14.8)
PHQ-8 depression ↓, mean (SD, range)	5.91 (4.99, 0-23)	5.94 (4.87, 0-23)	5.92 (4.96, 0-23)
General health ↓, mean (SD, range)	2.84 (0.760, 1-5)	2.87 (0.812, 1-5)	2.87 (0.770, 0-5)
Illness intrusiveness ↓, mean (SD, range)	2.63 (1.25, 1-7)	2.83 (1.11, 1-6.07)	2.80 (1.22, 1-7)
Hypoglycemic symptoms ↓, mean (SD, range)	1.35 (1.42, 0-6)	1.39 (1.40, 0-6)	1.39 (1.42, 0-6)
Fatigue ↓, mean (SD, range)	4.66 (2.31, 0-10)	4.86 (2.31, 0.10)	4.83 (2.31, 1-10)
Sleep ↓, mean (SD, range)	3.82 (2.89, 0-10)	3.92 (2.97, 0-10)	3.91 (2.91, 1-10)
Aerobic exercise (min/week) ↑, mean (SD, range)	95.3 (98.5, 0-555)	81.4 (104, 0-720)	84.0 (99.6, 0-720)
Stretching/range of motion (min/week) ↑, mean (SD, range)	34.0 (43.9, 0-180)	26.1 (49.8, 0-180)	27.6 (45.1, 0-180)
Communication with MD ↑, mean (SD, range)	2.78 (1.15, 0-5)	2.57 (1.23, 0-5)	2.61 (1.08, 0-5)
Medication adherence ↓, mean (SD, range)	0.971 (1.11, 0-4)	1.12 (1.02, 0-4)	1.09 (1.17, 0-4)
Proportion eye exam, last 12 months[b], %	131 (77.1)	533 (77.6)	664 (77.5)
Proportion foot exam, last 12 months[b], %	126 (74.1)	469 (69.7)	605 (70.5)
Proportion cholesterol exam, last 12 months[b], %	154 (90.6)	639 (93.0)	793 (92.5)
Proportion kidney exam, last 12 months[b], %	125 (73.5)	534 (77.7)	659 (76.8)

[a] Because of a lab recalibration changing HbA1c measurement, HbA1c prior to June 2014 was adjusted by adding 0.4. ↑ Indicates that a higher score is desirable, and ↓ that a lower score is desirable.

[b] "Don't know" set to no exam.

Figure 1. Face-to-face workshops CONSORT flowchart.

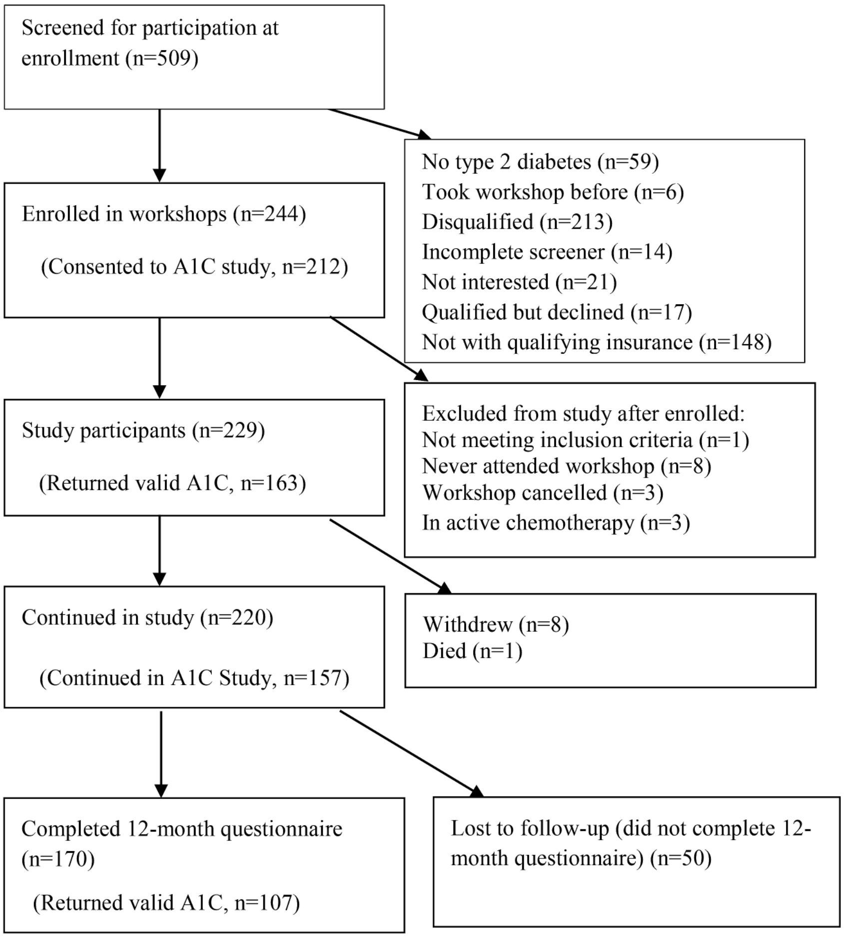
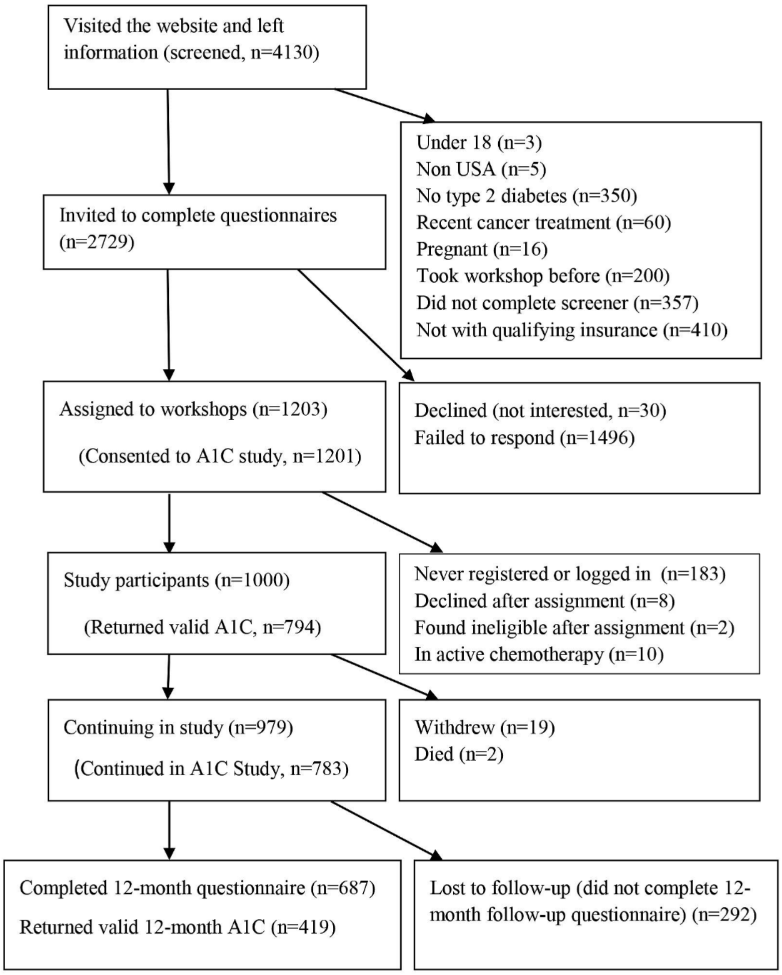


Figure 2. Web-based CONSORT flowchart.



Program Participation and Engagement

Participants took part in 1 of 50 face-to-face workshops or 49 Web-based workshops. Face-to-face workshops had 3-16 study participants attending a mean of 4.7 of 6 sessions (SD 1.5). Web-based workshops had 5-28 participants. Both modes also had nonstudy participants.

Web-based participants logged in a mean of 4.8 sessions (SD 1.5, range 1-6). They averaged 4.36 action plans (SD 3.9, range 0-6), 5.25 posts to discussion boards (3.9, 0-22), 10.4 replies to posts (20.1, 0-228), and 157.0 visits to workshop Web pages (62.4, 1-209).

Those Failing to Complete 12-Month Questionnaires

Of 1229 participants, 69.7% (857/1229) completed the 12-month questionnaire. The face-to-face participants (74.2%, 170/229) were more likely to complete compared to Web-based participants (68.7%, 687/1000). Outside of optional HbA1c, no variable had more than 1.7% missing data among 12-month completers. Nine-hundred and fifty-three consented to HbA1c testing and successfully completed baseline tests. Twelve-month HbA1c tests were successfully completed by 526 participants (55.0% (526/957) of those consenting and furnishing baseline HbA1c samples).

We compared baseline values of those not completing 12-month questionnaire and completers. Among demographic variables, noncompleters tended to be younger (mean 55.1 vs 57.8 years old, *P*<.001) and slightly less educated (14.9 vs 15.5 years of education, *P*<.001). There were no significant differences by gender, marital status, or ethnicity.

Among 15 outcome variables, 4 had statistically significance differences at baseline comparing completers and noncompleters. Also, 12-month noncompleters reported higher baseline HbA1c (8.4% vs 8.1% ,*P*=.007), lower general health (3.04 vs 2.87 on a 1-5 scale, *P*<.001), less aerobic exercise (65.3 vs 84.0 minutes/week, *P*=.002), and less likely to have had an eye exam in the last 12 months (79.7% vs 77.5%, *P*=.005).

Changes From Baseline

Table 2 shows the mean changes from baseline to 12 months (hypotheses 1). Seven of 7 health indicators (including HbA1c) and 7 of 8 health behaviors had statistically significant improvements between baseline and 12 months.

If we apply a Bonferroni correction and use .003 as the level of significance, 12 of 15 outcomes remain statistically significant.

Table 2. Baseline to 12-month scores, 12-month participants (N=857).

Variable[a]	Baseline mean (SD)	12-month mean (SD)	Baseline to 12-Month		*P* > |t| Baseline to 12 months
			Mean change	95% CI	
HbA1c ↓ [b] (n=526)	8.04 (1.44)	7.60 (1.55)	–0.447	–0.559 to –0.348	<.001
PHQ-8 depression (0-24) ↓	5.92 (4.96)	4.89 (4.87)	–1.02	–1.32 to –0.753	<.001
General health (0-5) ↓	2.87 (0.770)	2.77 (0.768)	–0.097	–0.140 to –0.048	<.001
Illness intrusiveness (1-7) ↓	2.79 (1.22)	2.65 (1.29)	–0.133	–0.214 to –0.065	<.001
Hypoglycemic symptoms (0-7) ↓	1.39 (1.42)	1.13 (1.29)	–0.260	–0.352 to –0.171	<.001
Fatigue (1-10) ↓	4.82 (2.31)	4.28 (2.52)	–0.541	–0.707 to –0.385	<.001
Sleep (1-10) ↓	3.91 (2.91)	3.68 (2.82)	–0.223	–0.427 to –0.040	.02
Aerobic exercise (min/week) ↑	84.0 (99.6)	101 (102)	16.7	9.91 to 23.4	<.001
Stretching or range of motion (min/week) ↑	27.6 (45.1)	34.0 (47.7)	6.28	2.83 to 9.52	<.001
Communication with MD (0-5) ↑	2.61 (1.08)	2.87 (1.20)	0.255	0.185 to 0.322	<.001
Medication adherence (0-4) ↓	1.09 (1.17)	0.917 (1.07)	–0.180	–0.240 to –0.103	<.001
Proportion eye exam, last 12 months (0,1) ↑[c]	0.778	0.849	0.070	0.035 to 0.105	<.001
Proportion foot exam, last 12 months (0,1) ↑[c]	0.710	0.799	0.090	0.056 to 0.123	<.001
Proportion cholesterol exam, last 12 months (0,1) ↑[c]	0.932	0.947	0.014	–0.007 to 0.036	.19
Proportion kidney exam, last 12 months (0,1) ↑[c]	0.776	0.885	0.109	0.077 to 0.141	<.001

[a] Possible range given in parentheses after variable name.

[b] Because of a lab recalibration changing HbA1c measurement, HbA1c prior to June 2014 was adjusted by adding 0.4. ↑ Indicates that a higher score is desirable, and ↓ that a lower score is desirable.

[c] "Don't know" set to no exam, change scores for those in both 6- and 12-month follow-ups.

12-Month Intent-to-Treat Analyses

For "intent-to-treat" analyses, we followed the standard practice of assuming no change for 12-month change scores for those who failed to complete 12-month questionnaires. The intent-to-treat analyses of changes resulted in no differences from the *P* values shown in Table 2.

HbA1c Changes

About 80% (957/1229) of participants supplied a valid HbA1c sample at baseline, and of these, 54.9% (526/957) supplied a 12-month HbA1c. There were no significant differences in any of the other baseline outcome measures for those who supplied an HbA1c sample versus those who did not. Participants who supplied an HbA1c sample had a slightly higher mean age (57 vs 55, *P*=.04), and were more likely to be of non-Hispanic white ethnicity (82.3% vs 68.5%, *P*<.001).

Because the lab that did the HbA1c testing recalibrated its measurements in June 2014 increasing all values by roughly 0.4%, all measures prior to June (some baseline and some 6-month) were adjusted upward by 0.4. If we remove all those with adjusted HbA1c at baseline (all whose baseline was before June 2014), the 12-month improvement in HbA1c remains almost the same (–0.454 vs –0.456). This suggests that our adjustment was appropriate.

Changes in Screening Tests (Hypotheses 1 and 2)

Table 3 gives the proportion of the participants who had each of 4 recommended tests in the year before baseline and the year between baseline and 12-month questionnaire. The proportion significantly increased for foot, eye, and kidney tests (*P*<.001). The proportion having cholesterol tests increased but not significantly, as it was already at a high level (93.2%) at baseline. For those not having a recommended test during the preintervention year, foot, eye, cholesterol, and kidney testing increased significantly. The total number of tests in the previous year also increased significantly from a mean of 3.2 (out of 4) to 3.5 (*P*<.001).

Table 3. Proportion receiving recommended examinations (N=768).

Type of exam[a]	Percent of those who had exam in 12 months prior to baseline, % (n)	Percent of those who had exam in 12 months between baseline and 1 year[b], % (n, *P* value)	Percent of those who had no exams 12 months prior to baseline against those who had exam in 12 months after baseline[c], % (n, *P* value)
Foot	71.0 (545)	80.0 (614, *P*<.001)	55.6 (124/223, *P*<.001)
Eye	77.9 (598)	84.9 (652, *P*<.001)	71.2 (121/170, *P*<.001)
Cholesterol	93.2 (716)	94.7 (742, *P*=.19)	76.9 (40/52, *P*<.001)
Kidney albumin	77.6 (592)	88.5 (680, *P*<.001)	72.7 (125/172, *P*<.001)

[a] Cases are those who had baseline, 6-month, and 12-month questionnaires.

[b] For *P* values in column 3, *t* tests were used to compare the proportion of those who had had the exam prior to baseline with the proportion of those who had had the exam between baseline and 12 months.

[c] For *P* values in column 4, chi-square test was used.

Web-Based Versus Face-to-Face Participants (Hypothesis 3)

Using *t* tests, only 4 1-year changes were found to be significantly different between Web-based and face-to-face participants: general health, sleep problems, stress, and medication adherence. These were further examined using general linear models controlling for the baseline values and for demographic covariates (age, gender, education, and whether minority). Differences in changes in general health were no longer significant after covariate adjustment. Web-based participants had greater improvements in medication adherence (*P*=.02), while face-to-face participants had greater improvements in sleep (*P*=.03) and stress (*P*=.002).

Older Versus Younger Participants (Hypothesis 4)

We segmented participants into those 65 and older (Medicare eligible, n=166) and those less than 65 years of age (n=691). There were 2 statistically significant differences in 12-month outcomes. The older group had significantly less increase in stretching and strengthening exercise (*P*=.01), and in general health (*P*=.02). For all other outcomes, change scores were similar. In particular, those over 65 decreased HbA1c by 0.43 while those under 65 similarly decreased HbA1c by 0.45.

Analyses of Participants With Specific Baseline Conditions (Hypothesis 5)

Because of sample heterogeneity, we examined change scores for specific outcomes for only those who had at least one of 5 criterial problems at baseline (high HbA1c, depression, hypoglycemia, low medication adherence, or no aerobic exercise).

High HbA1c

About 43% (524/1229) of the total sample had 12-month HbA1c results, and of those with 12-month HbA1c scores, 22.5% (118/524) had a baseline HbA1c of 9.0% or greater. By 12 months, approximately a third of these (30.3%, 36/119) had an HbA1c below 9. The mean decrease in HbA1c for those starting at above 9 was –1.27. Of those below 9 at baseline (n=407), the mean decrease was –0.206.

At baseline, 22.3% (117/524) had an HbA1c less than 7.0. By 12 months, 36% (189/524) of the participants were below 7, an increase of 61.5% (72/117). The 407 12-month participants with HbA1c of 7.0% or more at baseline had a mean 12-month decrease in HbA1c of 0.599, while those less than 7 at baseline (n=117) had a slight increase of 0.084.

No Aerobic Exercise

Participants who reported no aerobic exercise at baseline (n=185) had a mean increase of 46.0 minutes/week at 12 months. Those who reported some aerobic exercise had a mean increase of 8.6 minutes. The American Diabetes Association recommends that people with diabetes should have 150 minutes per week of moderate-intensity aerobic activity spread over 3 or more days [18]. At baseline, 22.3% (191/857) of the 12-month study sample met this criteria and this increased to 36.8% (316/857) at 12 months.

Low Medication Adherence

At baseline, 36.3% (311/856) reported low medication adherence. This decreased by 16% to 30.5% (261/856) by 12 months. Those with low adherence had a mean improvement (lower score) of –0.765. Adherent participants worsened by 0.154 (scale 0-4).

Hypoglycemic Symptoms

At baseline, 38.8% (329/849) of participants reported 2 or more hypoglycemic symptoms. At 12 months, 30.0% (255/849) reported 2 or more hypoglycemic symptoms. Those with symptoms at baseline had a mean improvement of –1.05. Participants without hypoglycemic symptoms worsened by 0.237 (scale 0-7).

Depression

At baseline, 21.3% (182/855) of participants had symptoms of depression, as defined by PHQ-8=10 or more [12]. At 12 months, the number with depression had decreased to 15.6% (133/855). The participants with depression had an improvement of –4.18 in PHQ-8 at 12 months. The participants with lower baseline PHQ-8 had a small improvement of –0.171 in PHQ-8 (scale 0-24).

Participants Who Improved in at Least One Condition (Hypothesis 6)

We examined the proportion of all participants (not limited to those with specific conditions) who improved in at least 1 of the 5 criterial conditions. Using effect-size improvements of at least 0.4 as an indication of an improvement, 79.9% of the total

12-month study population (685/857) improved in at least 1 of the 5 criterial variables. About 44% (377/857) improved in 2 or more. The mean number of improvements (out of a possible 5 for those with HbA1c data or 4 for those without) was 1.49 (SD 1.13). The 176 participants who did not improve in any of the 5 criterial variables had a mean of 0.956 criterial conditions, while those who improved in at least one criterion had a mean of 1.72 conditions. There was an r=0.374 correlation (Pearson) between the number of criterial conditions and the number of improvements of 0.4 effect size.

Discussion

Principal Findings

Between baseline (preintervention) and 1 year, there were modest, but statistically significant, improvements in 14 of 15 outcome measures, including HbA1c. There were no significant changes in the proportion having cholesterol exams within the preceding year. However, baseline percentages having those tests were high, leaving little room for improvement.

Healthy People 2020

The Healthy People 2020 [17] goals were to (1) reduce by 10% the proportion of persons with diabetes with an HbA1c value greater than 9.0%; (2) increase the proportion of the diabetic population with an HbA1c value less than 7% HbA1c by 10%; (3) increase the proportion of adults with diabetes who have at least an annual foot examination by 10% from a current 68% to 74.8%; (4) increase the proportion having annual eye exams by 10%; (5) increase the proportion having urinary microalbumin (kidney) tests by 10%; and (6) increase the proportion of adults with diabetes who receive formal diabetes education. By design, the BCBH-D workshops add to the proportion receiving diabetes education. For 4 of 5 goals, the participants in the program exceeded the desired increase of 10%. For the fifth goal (eye exams), the participants in the study were very close to achieving the goal (a mean increase of 9.2%). A year after the intervention, the participants exceeded many of Healthy People 2020 goals, such as improvements in HbA1c and percentage having recommended examinations. In addition, if the findings of this study can be replicated in other settings, the interventions may lead to improvement in both HEDIS and Medicare Star Quality Measures.

Limitations

The real-world nature of this pragmatic study necessitated the lack of a control group. Thus, we cannot be certain that the improvements observed are not due to other factors or might have occurred during the same time to nonparticipants. Alternative explanations for the improvements might include new medications and health plan initiatives that became available during the period of the study. It is also possible that there were interactions between workshop participation and members' likelihood of taking advantage of Anthem initiatives and new medications. While these other factors may be important, the consistency of the statistically significant improvements across multiple domains suggests a positive impact of workshop participation.

Attrition may also have affected outcomes (people with negative results may be more likely to avoid 12-month follow-up questionnaires). However, considering the large initial sample (n=1229) and large number of outcomes, there were only a few baseline statistical differences between 12-month completers and noncompleters, suggesting only minimal possible attrition bias.

For the analyses of changes for those with the 4 criteria conditions, regression to the mean might have contributed to the outcomes. However, the large differences in positive change scores for the worse-off compared with the smaller changes for the complementary subset of those relatively better-off suggest that not all of the improvements could have resulted from regression to the mean.

Even with the real-world pragmatic nature of the study, the participants were largely a self-selected population, perhaps more motivated than people who would not enroll. In 2016, the CMS Center for Innovation published a report from a 3000 plus person sample of people receiving Medicare [19]. Among other questions, they asked the likelihood of attending a wellness program if offered in their community in the next 6 months. Nearly 57% said that they were very likely (12.3%), likely (13.9%), or somewhat likely (30.6%) to enroll, and about half of these (46.4%) said they were interested in chronic diseases self-management programs.

Implications

As a community-based intervention, BCBH-D offered in 2 modes was associated with small but significant nonreinforced benefits, which were sustained for at least 1 year. The 2 different modes were able to reach somewhat different populations. This illustrates that offering more than 1 delivery mode reaches a broader population. Overall, there was little difference in 12-month improvements—participants in both delivery modes benefited similarly.

The program also appeared to meet some Healthy People 2020 diabetes objectives and could contribute to improved HEDIS and Medicare Star ratings. Most importantly, it was associated with clinically significant benefits for those with high HbA1c, for those with depression and hypoglycemia, as well as medication nonadherers and nonexercisers. The benefits differed by individual, but a large majority of the population had meaningful improvements in at least one of the areas. This study demonstrates that the peer-facilitated BCBH-D program, offered outside of the traditional health care system, can assist a national sample of health-plan members in improving their diabetes management, with benefits persisting for at least 1 year. This, and similar multidelivery-mode programs have the potential for making an impact on our growing diabetes population.

Acknowledgments

The Better Choices, Better Health-Diabetes self-management study received financial support from the Bristol Myers Squibb Foundation. The Anthem Foundation, the National Council on Aging (NCOA), National Council of Young Men's Christian Association of the United States of America (Y-USA), and OASIS Health aided in the administration and implementation of the program. Audrey Alonis and Maisoon Ayish at the Stanford Patient Education Research Center were responsible for data collection and management. Maurice Green was laboratory consultant for the HbA1c testing. Kathlyn Plant, then with NCOA, helped with recruitment and was liaison between Stanford and NCOA researchers for the online program. James Teufel, of OASIS Health, assisted with recruitment and IRB submission. Sarah Lovegreen, OASIS Health, and Valerie Lawson, Y-USA, helped with recruitment and administration of community programs. Julie Kosteas, Kristie Kulinski, Simona Eldridge, NCOA, assisted with administration of the online program, and Andrea DeVries, HealthCore, assisted with reviewing analyses and presentation.

Conflicts of Interest

Kate Lorig and Diana D Laurent have the potential to receive royalties if the program is further disseminated and receive royalties from the book used in the intervention.

References

1. Li R, Shrestha SS, Lipman R, Burrows NR, Kolb LE, Rutledge S, Centers for Disease Control and Prevention (CDC). Diabetes self-management education and training among privately insured persons with newly diagnosed diabetes--United States, 2011-2012. MMWR Morb Mortal Wkly Rep 2014 Nov 21;63(46):1045-1049 [FREE Full text] [Medline: 25412060]
2. NCQA. HEDIS 2015. 2016. URL: http://www.ncqa.org/hedis-quality-measurement/hedis-measures/hedis-2015 [accessed 2016-08-10] [WebCite Cache ID 6jfNStQGp]
3. Lorig K, Ritter PL, Turner RM, English K, Laurent DD, Greenberg J. Benefits of diabetes self-management for health plan members: a 6-month translation study. J Med Internet Res 2016;18(6):e164 [FREE Full text] [doi: 10.2196/jmir.5568] [Medline: 27342265]
4. American Diabetes Association. Standards of medical care in diabetes-2015 abridged for primary care providers. Clin Diabetes 2015 Apr;33(2):97-111. [doi: 10.2337/diaclin.33.2.97] [Medline: 25897193]
5. Lorig K, Ritter PL, Laurent DD, Plant K, Green M, Jernigan VB, et al. Online diabetes self-management program: a randomized study. Diabetes Care 2010 Jun;33(6):1275-1281 [FREE Full text] [doi: 10.2337/dc09-2153] [Medline: 20299481]
6. Lorig K, Ritter PL, Villa FJ, Armas J. Community-based peer-led diabetes self-management: a randomized trial. Diabetes Educ 2009;35(4):641-651. [doi: 10.1177/0145721709335006] [Medline: 19407333]
7. Bandura A. Self-efficacy: the exercise of control. New York: WH Freeman; 1997.
8. Coremedicalabs. Dried blood spot (DBS). URL: https://www.coremedicalabs.com/blood-collection.html [accessed 2016-01-25] [WebCite Cache ID 6epZSghvN]
9. Stanford University. Research instruments developed, adapted or used by the Stanford Patient Education Research Center. URL: http://patienteducation.stanford.edu/research/ [accessed 2016-08-09] [WebCite Cache ID 6jfMxmxN2]
10. Lorig K, Stewart A, Ritter P, González V, Laurent D, Lynch J. Outcome measures for health education and other health care interventions. Thousand Oaks, CA: Sage Publications; 1996.
11. Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. Am J Public Health 1990 Apr;80(4):446-452. [Medline: 2316767]
12. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
13. Devins GM, Mandin H, Hons RB, Burgess ED, Klassen J, Taub K, et al. Illness intrusiveness and quality of life in end-stage renal disease: comparison and stability across treatment modalities. Health Psychol 1990;9(2):117-142. [Medline: 2331973]
14. Piette JD. Patient education via automated calls: a study of English and Spanish speakers with diabetes. Am J Prev Med 1999 Aug;17(2):138-141. [Medline: 10490057]
15. Ritter PL, González VM, Laurent DD, Lorig KR. Measurement of pain using the visual numeric scale. J Rheumatol 2006 Mar;33(3):574-580. [Medline: 16511926]
16. Morisky DE, Green LW, Levine DM. Concurrent and predictive validity of a self-reported measure of medication adherence. Med Care 1986 Jan;24(1):67-74. [Medline: 3945130]
17. Colberg SR, Sigal RJ, Fernhall B, Regensteiner JG, Blissmer BJ, Rubin RR, American College of Sports Medicine, American Diabetes Association. Exercise and type 2 diabetes: the American College of Sports Medicine and the American Diabetes Association: joint position statement executive summary. Diabetes Care 2010 Dec;33(12):2692-2696 [FREE Full text] [doi: 10.2337/dc10-1548] [Medline: 21115771]
18. Healthypeople. Diabetes. 2016. URL: http://www.healthypeople.gov/2020/topics-objectives/topic/diabetes [accessed 2016-01-24] [WebCite Cache ID 6epYzLSbH]
19. Centers for Medicare & Medicaid Services. Wellness Prospective Evaluation Report on Baseline Survey Efforts and Qualitative Study of Program Operations and Costs. URL: https://innovation.cms.gov/Files/reports/communitywellnessprgms-frstevalrpt.pdf [accessed 2016-12-05] [WebCite Cache ID 6mXF73zHv]

Abbreviations

BCBH-D: Better Choices, Better Health-Diabetes
CLIA: Clinical Laboratory Improvement Amendments
HEDIS: Healthcare Effectiveness Data and Information Set
PHQ-8: Patient Health Questioinaire

Edited by G Eysenbach; submitted 10.08.16; peer-reviewed by K Blondon, E Fisher; comments to author 15.09.16; revised version received 17.10.16; accepted 23.11.16; published 15.12.16.

Please cite as:
Lorig K, Ritter PL, Turner RM, English K, Laurent DD, Greenberg J
A Diabetes Self-Management Program: 12-Month Outcome Sustainability From a Nonreinforced Pragmatic Trial
J Med Internet Res 2016;18(12):e322
URL: http://www.jmir.org/2016/12/e322/
doi:10.2196/jmir.6484
PMID:27979790

Original Paper

Drivers and Barriers to Acceptance of Web-Based Aftercare of Patients in Inpatient Routine Care: A Cross-Sectional Survey

Severin Hennemann[1,2], Dipl-Psych; Manfred E Beutel[1], PhD; Rüdiger Zwerenz[1], PhD

[1]Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Gutenberg University Mainz, Mainz, Germany

[2]Department of Clinical Psychology, Psychotherapy and Experimental Psychopathology, Institute of Psychology, University of Mainz, Mainz, Germany

Corresponding Author:
Rüdiger Zwerenz, PhD
Department of Psychosomatic Medicine and Psychotherapy
University Medical Center
Gutenberg University Mainz
Untere Zahlbacher Str. 8
Mainz, 55131
Germany
Phone: 49 6131 175981
Fax: 49 6131 175563
Email: ruediger.zwerenz@unimedizin-mainz.de

Abstract

Background: Web-based aftercare can help to stabilize treatment effects and support transition after inpatient treatment, yet uptake by patients seems limited in routine care and little is known about the mechanisms of adoption and implementation.

Objective: The aim of this study was to (1) determine acceptance of Web-based aftercare and (2) explore its drivers and barriers in different subgroups of a mixed inpatient sample.

Method: In a cross-sectional design, 38.3% (374/977) of the inpatients from a broad spectrum of diagnostic groups (psychosomatic, cardiologic, orthopedic, pediatric, and substance-related disorders) filled out a self-administered questionnaire prior to discharge. Drivers and barriers to patients' acceptance of Web-based aftercare were examined based on an extension to the "unified theory of acceptance and use of technology" (UTAUT). In total, 16.7% (59/353) of the participants indicated prior use of eHealth interventions.

Results: Acceptance (min 1, max 5) was low (mean 2.56, SD 1.22) and differed between diagnostic groups (Welch $F_{4,133.10}=7.77$, $P<.001$), with highest acceptance in adolescent patients (mean 3.46, SD 1.42). Acceptance was significantly predicted by 3 UTAUT predictors: social influence (beta=.39, $P<.001$), performance expectancy (beta=.31, $P<.001$), and effort expectancy (beta=.22, $P<.001$). Furthermore, stress due to permanent availability (beta=−.09, $P=.01$) was negatively associated with acceptance.

Conclusion: This study demonstrated a limited acceptance of Web-based aftercare in inpatients. Expectations, social environment's attitude, and negative experience with permanent availability influence eHealth acceptance. Improving implementation, therefore, means increasing eHealth experience and literacy and facilitating positive attitudes in patients and health professionals through education and reduction of misconceptions about effectiveness or usability.

(J Med Internet Res 2016;18(12):e337) doi:10.2196/jmir.6003

KEYWORDS

eHealth; attitude to health; aftercare; rehabilitation; survey

Introduction

Background

The digitalization of health care brought forth a broad range of effective eHealth interventions for various somatic [1] and mental health problems [2]. The effects of guided interventions for some mental disorders turned out to be comparable to those of traditional face-to-face therapy [3]. eHealth interventions can be a cost-efficient, time-flexible, and location-independent extension of existing health care [4], particularly in areas such as prevention, as aftercare or blended care. In the case of inpatient rehabilitation, previous research focused on Web-based aftercare [5], as the transition to daily life after inpatient treatment is a critical phase, with elevated risk of relapse, sick leave, return to adverse health behavior, or even inpatient readmission [6]. Therefore, clinics often recommend participation in outpatient treatment and aftercare (eg, physiotherapy, vocational rehabilitation, educative group sessions). However, organizational constraints or limited accessibility, concerns about anonymity, or negative treatment expectancies can reduce the utilization of conventional outpatient treatment or aftercare [7,8]. In this regard, Köpke [9] describes insufficient uptake rates of conventional aftercare interventions between 18.9% for somatic and 69.6% for psychosomatic disorders after inpatient rehabilitation in Germany.

Previous Web-based aftercare interventions have been developed and evaluated particularly for psychosomatic inpatients, for example, cognitive behavioral therapy (CBT)–based Internet platforms [10,11] or mobile interventions [12], providing modular or consecutive treatment elements with a combination of health-related information and interactive tasks to promote health behavior (see also Multimedia Appendix 1). Furthermore, interventions differ by amount of guidance by health experts [13], who support health behavior and self-efficacy through synchronous (eg, video- or chat-based) [14] or asynchronous (email, short message service) [11] interaction and task-related feedback. Most Web-based aftercare interventions with psychosomatic focus are CBT based, whereas more recent developments include psychodynamic [15] or acceptance- and commitment-based approaches [16].

However, despite an increased proportion of inpatients with pronounced work-related stress and subsequently an elevated risk of early retirement [17], occupational aspects play a subsidiary role in eHealth interventions in the context of inpatient treatment. Only a few projects explicitly addressed work-related stress in clinical or subclinical samples and yielded mixed effects [15,18,19]. As the digitalization of health care is still in its early stages in Germany, eHealth interventions are scarce in inpatient care, with only 9% of studies on eHealth interventions in the area of rehabilitation according to Eichenberg and Ott [20].

A fundamental precondition to the implementation of eHealth interventions is the willingness to utilize and adopt new technologies by patients and help seekers [21]. Despite their efficacy, uptake of and adherence in eHealth interventions pose a challenge for implementation in routine care [22], and mechanisms of uptake of eHealth interventions and their relevance to implementation have not been studied sufficiently yet.

Acceptance of E-Mental Health Interventions

Acceptance can imply the utilization of an intervention, as expressed in uptake rates, adherence, or satisfaction. Following theoretical models of acceptance of information technology, acceptance can also be defined as the behavioral intention to use eHealth interventions [23,24] as an essential prerequisite to adoption and implementation of eHealth interventions.

Based on their review of 36 studies, Waller and Gilbody [25] concluded that participants' satisfaction with computerized CBT (cCBT) is usually high; however, the willingness to participate is limited, with only 38% starting the intervention after recruitment. Other findings are inconsistent, with various studies documenting a low uptake rate of e-mental health interventions by patients [21,25] or limited acceptance in the general population [26,27].

According to a recent representative population survey in Germany, less than 10% of the interviewed participants could imagine using psychological Web-based support and merely 2% had already used such services in the past [27]. Whereas recent studies point to the acceptance of eHealth records and computer-based diagnostic systems in inpatient treatment [28,29], preliminary evidence on some Web-based aftercare projects in German rehabilitation shows fluctuating uptake rates between 21% and 62% [30,12]. Multimedia Appendix 1 compares acceptance ratings (eg, uptake, adherence) of exemplary Web-based aftercare interventions including guided self-help [11], blended [31], mobile-based [12,32], and chat- or education based [33-36] interventions. The results of Pfaudler et al [33] and Küffner [35] indicate organizational and technical problems as pivotal barriers to adoption by inpatients.

Methodical Shortcomings of Acceptance Assessment

Although acceptance ratings have been included in most studies on eHealth interventions by now, methodical shortcomings may somehow limit their validity. Since acceptance is typically assessed retrospectively, no information about genuine attitudes toward this form of health promotion or reasons for intention to use an intervention is provided. Furthermore, acceptance ratings may be confounded with intervention satisfaction and do not reflect eHealth acceptance or attitudes. A selective dropout of dissatisfied participants and participation of Internet-oriented patients may also positively bias acceptance ratings. The lack of sufficient theoretical foundation and valid or consistent instruments of acceptance evaluation can further be criticized [21]. A recent example for a methodical valuable approach is the 4-dimensional "Attitudes toward psychological online interventions" questionnaire (APOI), which has been cross-validated in a large sample of individuals with symptoms of depression [37]. Furthermore, previous studies focused on acceptance in rather selected patient samples. Hence, comprehensive findings from inpatient treatment are scarce. Altogether, these methodological shortcomings may illustrate the importance of a differentiated, theory-based, and transdiagnostic assessment of eHealth acceptance and its preconditions.

Barriers and Resources to eHealth Acceptance

Research on predictors of eHealth acceptance has established the technology acceptance model (TAM) [23], and its advancement, the unified theory of acceptance and use of technology (UTAUT) [24] in several studies in patient populations [38-41] or health professionals [42-44]. Both models integrate traditional motivational theories such as the theory of planned behavior [45] or Bandura's social cognitive theory [46]. Since the UTAUT is more differentiated regarding predictors on acceptance and has recently been used in similar research questions [38-41], we adapted this model to the following study. The UTAUT postulates 4 core predictors of behavioral intention: performance expectancy (expected benefit of technology use), effort expectancy (expected ease of use), social influence (attitude of significant others toward using the technology), and facilitating conditions (organizational or technical resources and preconditions necessary to technology use). In previous studies, performance expectancy proved to be the most important predictor of eHealth acceptance [44,47,48]. In a meta-analysis of 37 studies testing the UTAUT in various contexts, Taiwo and Downe [48] found a medium aggregated effect size of correlation with acceptance (z-transformed r=.54) and somewhat smaller effect sizes for the remaining model predictors (.38-.44).

However, research also points out the complex nature of acceptance and its determinants [42,43]. Therefore, the predictive model of eHealth acceptance needs to be extended and adapted to the context of different target groups. Typically, sociodemographic characteristics such as younger age, higher educational level, Internet access, and technical experience have been associated with acceptance of eHealth interventions in several studies [2,27,49,50]. Further potential barriers to acceptance include Internet anxiety [40,51], low Internet orientation in health problems [27], insufficient knowledge of eHealth interventions [27,52], rural residence [53], or reservations regarding data security or impersonal interaction [54,55]. However, barriers or facilitators to eHealth adoption in inpatient routine care have not been studied sufficiently yet.

Goals of This Study

This study aimed to (1) determine the current status of acceptance of Web-based aftercare with a focus on work-related stress and (2) explore its drivers and barriers based on an extended UTAUT model in a mixed sample of inpatients in a cross-sectional survey. Besides previously studied Internet-related predictors, we were interested in the influence of eHealth literacy, which is defined as the ability to find, evaluate, and utilize Internet-based health information to deal with health problems [56]. Furthermore, we aimed to extend existing literature on the role of self-efficacy [57] in the eHealth context. Since the evidence on the influence of symptom severity and uptake of e-mental health interventions is mixed [11,58], we aimed to explore the relation of mental and occupational distress with acceptance. Also, since modern technologies increase digital load and availability [59], we were interested in the impact of stress due to constant availability (email, social media) on acceptance.

We postulated positive relations between UTAUT- and Internet-related predictors as well as self-efficacy. We assumed that Internet anxiety and stress through permanent availability would be negatively associated with acceptance ratings. Furthermore, we investigated subgroup-specific differences. Here, we expected younger age, higher educational status, and urban residency to be associated with higher acceptance.

Methods

Study Design

A cross-sectional approach was adopted to compare acceptance and its predictors in different inpatient diagnostic groups with a self-administered questionnaire. The survey was based on a qualitative pilot study, in which inpatients of various diagnostic groups were interviewed on health-related Internet use and barriers or resources to eHealth utilization in semistructured interviews. The questionnaires were built upon previous findings on predictors of acceptance, which were adapted to this study. In a consensual approach, the content, clarity, and face validity of the items were appraised by internal and external researchers and pretested in one of the cooperation clinics. The survey was conducted in 4 inpatient rehabilitation centers of the German statutory pension insurance scheme, covering a broad range of the most common diagnostic groups: psychosomatic medicine and psycho-oncology (PSY), orthopedics (OPE), cardiology (CAR), pediatric disorders of adolescent patients (PED), and substance use disorders (SUD).

Participants and Recruitment

Participants were recruited consecutively by 4 research assistants from August 2015 to January 2016. Participants were required to be aged 14 or above due to ethical rules. Inpatients with at least two weeks of treatment or half of treatment completed were included in the study. Since the focus on work-related stress in Web-based aftercare, retired patients were excluded from further analyses. Of 977 eligible patients who received a questionnaire, 374 (38.3%) responded. Participants were first informed at admission (eg, in an introductory course or at intake) and recruited subsequent to a regular group screening prior to discharge. After oral and written information, patients could fill out the paper-pencil questionnaire. The survey contained 69 questions (59 in pediatric sample, as items regarding occupational aspects were omitted) and took about 15 minutes to fill out. The survey was completely anonymous. No written consent was needed for participants older than 15 years. However, 14-year-old participants and their parents had to fill out a written consent. Gift cards served as an incentive for participation. Participants who wanted to take part in the draw of the gift cards could leave their contact details, which were detached from the questionnaire and kept separately and inaccessible to the project manager to ensure anonymity. All procedures involved in the study were approved by the ethics committee of the Federal State of Rhineland Palatinate, Germany (Ref. No. 9434-F) and the data protection commissioner of the German statutory pension insurance scheme.

Measures

Primary Outcome: Acceptance of Web-Based Aftercare

Based on the UTAUT [23], acceptance was operationalized as the intention to use Web-based aftercare. Acceptance was measured using 3 items (Table 1). All items were rated on 5-point Likert scales ranging from (1) “totally disagree” to (5) “totally agree,” with higher scores indicating elevated acceptance. The items were adapted from previous studies [23,38-40,47]. Following a short introduction with general information about Web-based aftercare, the prefix to all items was “A Web-based aftercare for the management of work-related stress...” (for adolescent patients: “Aftercare delivered over the Internet...”). Reliability was calculated as Cronbach alpha. To avoid underestimation of true reliability, the internal consistency of 2-item subscales was not calculated [60]. Cronbach alpha of the acceptance scale in this study was .85.

Table 1. Adapted items of the UTAUT model and references of original studies.

Variable	Items
Behavioral Intention	“I would like to try a Web-based aftercare.”[a,b,c]
	“I would use a Web-based aftercare if offered to me.”[a,b,c]
	“A Web-based aftercare would be worth paying for.”[a,b,c]
Social influence	“People close to me would approve a Web-based aftercare.”[a,b,c,d]
	“My general practitioner would approve the use of a Web-based aftercare.”[a,b,c]
	“My friends would approve a Web-based aftercare.”[e]
Performance expectancy	“A Web-based aftercare could improve my work-related well-being.”[a,b,c]
	“A Web-based aftercare could help me with work-related stress.”[a,b,c]
	“A Web-based aftercare could help me to improve my personal health.”[a,b,c,e]
	“A Web-based aftercare could help me with my health problems.”[a,b,c,e]
Effort expectancy	“A Web-based aftercare would be easy to operate and comprehend.”[a,b,c,d,f]
	“I could arrange using a Web-based aftercare in my everyday life.”[g]
Facilitating conditions	“I have the necessary technical preconditions for using a Web-based aftercare.”[a,b,c,d]
	“I possess the required technical knowledge to utilize a Web-based aftercare.”[c,d,f]

[a] Baumeister et al [38].

[b] Baumeister et al [39].

[c] Ebert et al [40].

[d] Venkatesh et al [23].

[e] Items used in the adolescent sample (PED).

[f] Liu et al [47].

[g] Self-constructed.

Secondary Outcome: Predictors of Acceptance

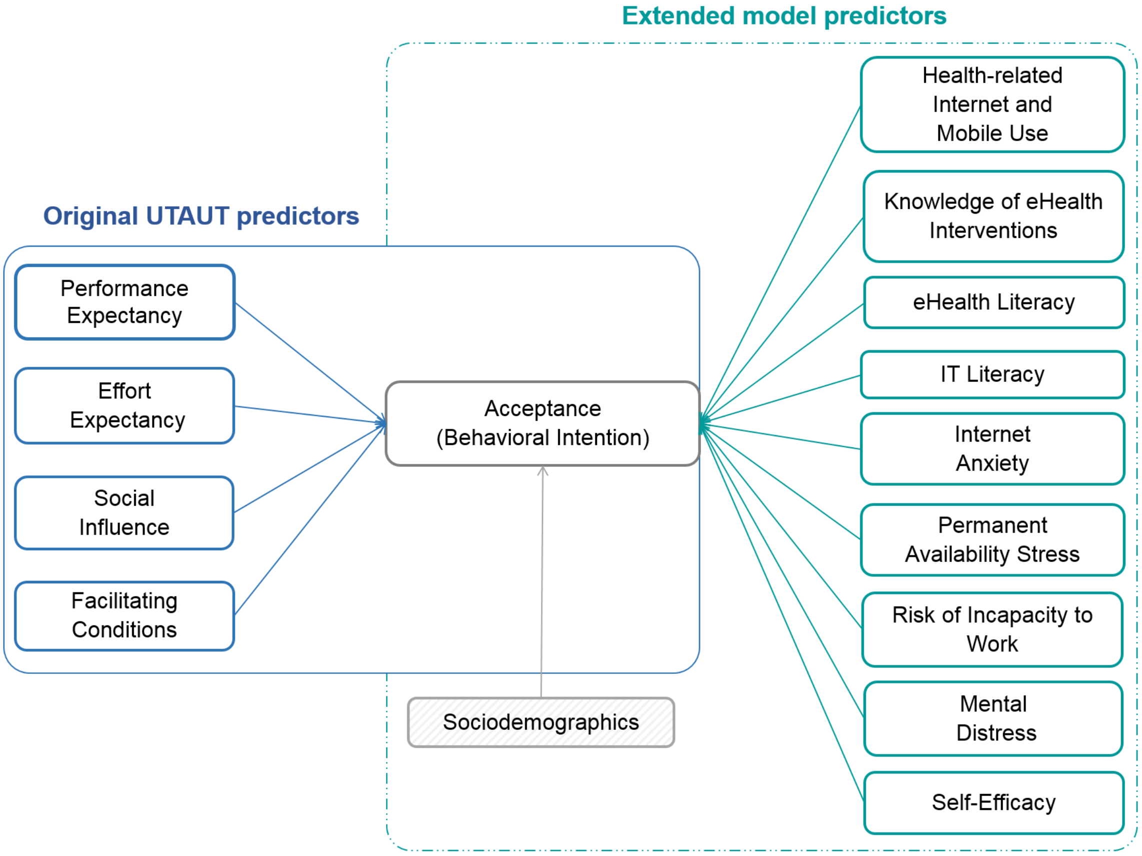
The UTAUT predictors were measured using 2 items each, adapted from previous studies [23,38-40,47]. Answers were rated on a 5-point Likert scale ranging from from (1) “totally disagree” to (5) “totally agree.” The items of the UTAUT predictors used in the questionnaire and the original studies they were adapted from are listed in Table 1.

Further differential factors were included: *knowledge of eHealth interventions* was measured in 2 items following a short explanation of eHealth interventions based on a previous study by Ebert et al [40]. Two items measuring *Internet anxiety* were adapted from previous studies [38-40]. Answers were rated analogous on a 5-point Likert scale ranging from (1) “totally disagree” to (5) “totally agree.” The frequency (from [1] “never” to [5] “always”) of *health-related Internet and mobile use* was evaluated for different applications (fitness, vital parameters, nutrition, mental and occupational problems). *Self-efficacy* was assessed with the German general self-efficacy short form (ASKU) [61], consisting of 3 items with responses rated on a 5-point Likert-type scale ranging from (1) “does not apply at all” to (5) “applies completely.” The questionnaire has proven to be reliable with alpha=.81 to .86 [62]. Cronbach alpha in this study was .89.

eHealth literacy was measured with the German adaption of the eHealth literacy scale (eHEALS) [63], consisting of 8 items. Answers were rated on a 5-point Likert scale ranging from (1) “strongly disagree” to (5) “strongly agree.” Previous research by Norman et al [64] showed a high internal consistency of .88. For economic reasons, the questionnaire for adolescent patients included only 3 items of the eHEALS. Items were selected regarding linguistic appropriateness for the target population and high factorial loadings [63]. Cronbach alpha of the original scale in this study was .95.

Mental distress was assessed with the Patient Health Questionnaire-4 (PHQ-4) as a brief screening instrument for depression and anxiety [65] in the adult patient sample. In 4 items with answers on a 4-point scale ranging from 0 (“Not at all) to 3 (“Nearly every day”), the extent of symptoms in the last 2 weeks was evaluated. Sum scores can be categorized as normal (0-2), mild (3-5), moderate (6-8), and severe (9-12) [65]. Cronbach alpha in this study was .93.

Figure 1. Research model based on the UTAUT (left) and extended predictors (right). UTAUT: unified theory of acceptance and use of technology.



Furthermore, we measured the *subjective risk of incapacity to work* with the 3-item subjective prognosis of work ability scale (SPE) [66] in adult participants. Answers were dichotomized, and sum scores ranged from 0 to 3. Probability of retirement doubles with each scale interval [67]. Cronbach alpha in this study was .70. Additionally, the influence of *IT-literacy* [answers ranging from (1) “very little knowledge” to (5) “very good knowledge”] and *stress through permanent availability* (level of agreement on a 5-point Likert-type scale from [1] “strongly disagree” to [5] “strongly agree”) on acceptance were examined in self-constructed items.

To further explore preconditions of intended usage, participants could state subjective advantages or disadvantages of using a Web-based aftercare. Finally, 1 item measured preferences for different forms of Internet-based support. The research model is depicted in Figure 1.

Sociodemographic Characteristics

The survey contained items regarding sex, age, education or current school type, population of residence, Internet access, time spent on Internet in free time per day, occupational status prior to rehabilitation, and days on sick leave (from [1] “less than 9 days” to [5] “100-365 days”) from the Work Ability Index [68]. One item differentiating between types of experience with eHealth interventions was adapted from a previous study by Eichenberg et al [27].

Statistical Analyses

The data were analyzed using SPSS Statistics version 23 (IBM). Out of 374 datasets, 18 individuals with missing values exceeding 50% (mean 77.73% missing, SD 17.37) were excluded from further analysis [69]. Furthermore, 18 individuals who indicated being retired were excluded, resulting in 338 participants analyzed. We addressed missing data under the assumption of data missing at random with multiple imputation technique in SPSS including all analysis variables [62]. To demonstrate the extent of missing data, sociodemographic statistics (Multimedia Appendix 2) were calculated with observed data, whereas further analyses were based on imputed values (0.6%-15.7% missing values per variable). Imputed statistics were comparable to observed statistics. Group differences of demographics, primary, and secondary outcomes were analyzed using the chi-square test or G-test, analysis of variance with post-hoc comparisons (Games-Howell), or independent *t* tests. In case of variance inhomogeneity, Welch F-ratio was calculated. Acceptance scores were further categorized by mean to describe low (1-2.34), moderate (2.35-3.67), and high (3.68-5) acceptance, and frequency of categories was calculated. The predictive model of acceptance was tested using multiple hierarchical regression. Predictors were included blockwise: (1) sociodemographic and background variables (mental and occupational distress, self-efficacy), (2) Internet-related variables, and (3) UTAUT predictors. According to Midi, Sarkar, and Rana [70], no sign of severe multicollinearity could be detected and all predictors could enter the regression. To perform further between-group comparisons of acceptance, mental distress and risk of incapacity to work were dichotomized with a median split. Open answers regarding advantages and disadvantages of Web-based aftercare were rated according to qualitative content analysis [71] by 2 independent raters following a-priori consensually developed categories. An inter-rater reliability analysis using Cohen kappa statistic was performed to evaluate consistency.

Results

Sociodemographic Characteristics

General characteristics of 374 participants are shown in Multimedia Appendix 2. More men than women participated in the study, and mean age was 45.49 years (SD 14.17, min 14, max 78). The age range of the pediatric subsample was 14-18 years. The majority of adult inpatients indicated being employed prior to treatment. Educational status was moderate overall. A total of 59 (16.7%, 59/353) participants stated prior eHealth use. Significant differences between diagnostic groups were found in all sociodemographic variables with exception to the distribution of wearables and prior e-mental health usage, primarily in comparison to the pediatric subsample.

Acceptance of Web-Based Aftercare

Acceptance of Web-based aftercare was overall low in patients with a mean of 2.56 (SD 1.23). The highest acceptance was found in pediatric patients with a mean of 3.46 (SD 1.42), the lowest in orthopedic patients with a mean of 2.18 (SD 1.00). The diagnostic groups differed statistically significantly (Welch $F_{4,133.10}$=7.77, P<.001). Post-hoc comparisons demonstrated significant differences in all disorder groups compared with the pediatric patients, but not between other diagnostic groups. When including sociodemographic variables as covariates, diagnostic group differences remained marginally significant ($F_{4,\ 326}$=2.41, P=.05). Acceptance between diagnostic groups is depicted in Figure 2.

Further analyses revealed that 48.8% (165/338) could be categorized with a low, 32.8% (111/338) with a moderate, and 18.4% (62/338) with a high acceptance. This pattern only differed in cardiological (42% [28/67] “low,” 48% [32/67] “moderate,” 10% [7/67] “high”) and pediatric patients (22% [11/51] “low,” 24% [12/51] “moderate,” 55% [28/51] “high”). Only 10.2% (31/305) participants with valid answers agreed that they would be willing to pay for participation in a Web-based aftercare.

Acceptance differed significantly between age groups ($F_{3,342}$=9.97, P<.001) with highest acceptance in the youngest quartile (14-26 years). Post-hoc tests revealed significant differences between the youngest and all other age groups. Further data inspection of sociodemographic variables revealed no differences in acceptance regarding gender, population of residence, and occupational status. However, prior eHealth use, higher educational status, and private Internet access were associated with higher acceptance. Table 2 contains acceptance scores as a function of demographics.

Figure 2. Acceptance (means) of Web-based aftercare between diagnostic groups. Sample size in brackets. Error bars represent standard deviations.

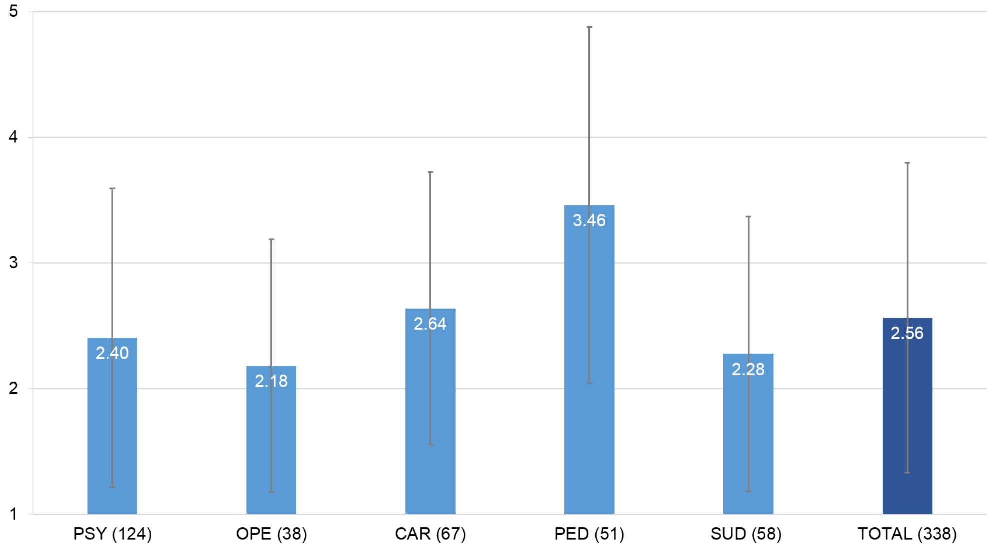


Table 2. Differences in acceptance by demographics (N=338).

Variable	n	Mean (SD)	Test	*P* value
Sex				
Male	198	2.62 (1.21)	t_{336}=0.93	.35
Female	140	2.49 (1.26)		
Age in years				
14-26	57	3.34 (1.42)	$F_{3,334}$=9.57	*P*<.001
27-39	32	2.40 (1.18)		
40-52	130	2.41 (1.11)		
53-65	119	2.40 (1.23)		
Educational status				
No graduation	10	1.95 (1.04)	$F_{2,335}$=8.00	*P*<.001
Secondary school	277	2.48 (1.18)		
A-level	51	3.14 (1.38)		
eHealth experience				
Yes	56	2.44 (1.20)	Mann-Whitney *U*=5315	*P*<.001
No	282	3.16 (1.24)		
Internet access				
Yes	296	2.65 (1.23)	t_{336}=−3.60	*P*<.001
No	42	1.98 (1.11)		
Occupational status (287)[a]				
Employed	195	2.43 (1.13)	$F_{2,284}$=0.41	.66
Unemployed	44	2.26 (1.11)		
Sick leave	48	2.42 (1.14)		
Population of residence				
Rural	145	2.53 (1.27)	t_{336}=−0.41	.68
Urban	193	2.59 (1.21)		

[a] Retired and adolescent patients (PED) excluded.

Drivers and Barriers to Acceptance

Multiple hierarchical regression in adult inpatients (PSY, OPE, CAR, SUD) showed that out of the demographic and background variables entered in the first step, gender (beta=–0.13, *P*=.04) and Internet access (beta=0.13, *P*=.03) predicted acceptance, although explained variance was low (R^2=.06, $F_{9,272}$=2.09, *P*=.03). When entering Internet-related variables, explained variance increased significantly (R^2=.17, $F_{17,264}$=3.22, *P*<.001), but only health-related Internet and mobile use significantly predicted acceptance (beta=0.22, *P*<.001). The full regressive model including UTAUT variables showed high explained variance (R^2=.78, $F_{21,260}$=43.86, *P*<.001). Social influence (beta=.39, *P*<.001), performance expectancy (beta=.31, *P*<.001), and effort expectancy (beta=.22 , *P*<.001) significantly predicted acceptance, whereas facilitating conditions did not reach significance (Table 3). Furthermore, stress through permanent availability (beta=–.09, *P*=.01) could be observed as a fourth significant predictor. Table 3 contains the regression parameters of the complete regression model.

In a separate regression analysis of the adolescent subsample (PED), explained variance of acceptance significantly improved in the model including UTAUT-predictors (R^2 = .85, $F_{18,32}$=10.28, *P*<.001) in comparison to the previous steps (demographic variables only: R^2 = .24, $F_{6,44}$=2.33, *P*=.05; demographic and Internet-related variables: R^2=.43, $F_{14,36}$=1.94, *P*=.06). However, of the UTAUT-predictors, only performance expectancy (beta=.67, *P*<.001) significantly predicted acceptance of Web-based aftercare. Knowledge of eHealth interventions (beta=.21, *P*=.02) and IT literacy (beta=−.22, *P*=.04) proved to be further predictors of acceptance in the full model. The correlation matrix of the total sample can be found in the Multimedia Appendix 3. Absolute values of intercorrelations ranged from |0.0| to |0.79|.

Table 3. Regression model of acceptance (full model, adult sample, N=282).

Predictor	b	Standard error	95% CI	Beta	*P* value
Constant	.08	.44	−0.78 to 0.94		.86
Sociodemographic predictors					
Sex	.03	.07	−0.11 to 0.18	.01	.64
Age	.01	.00	0.00 to 0.02	.05	.17
Days on sick leave	–.01	.03	−.0.06 to 0.05	−.01	.76
Internet access	–.05	.11	−0.27 to 0.18	−.01	.68
Population of residence	–.11	.07	−0.25 to 0.03	−.05	.12
Educational status	–.06	.09	−0.24 to 0.11	−.02	.49
Self-efficacy[a]	–.01	.05	−0.10 to 0.08	−.01	.80
Mental distress[b]	.01	.01	−0.02 to 0.03	.02	.66
Risk of incapacity to work[c]	–.01	.03	−0.07 to 0.06	−.01	.85
Internet-related predictors					
eHealth literacy[d]	.01	.06	−0.10 to 0.12	.01	.90
Internet anxiety	.02	.04	−0.06 to 0.10	.02	.67
Knowledge of eHealth interventions	–.06	.04	−0.14 to 0.02	−.06	.14
Time on Internet	–.01	.02	−0.04 to 0.03	−.01	.71
eHealth experience	.18	.11	−0.03 to 0.39	.05	.09
Health-related Internet and mobile use	.07	.06	−0.04 to 0.19	.04	.21
IT literacy	.00	.04	−0.09 to 0.08	.00	.95
Permanent availability stress	–.08	.03	−0.14 to −0.02	−.09	.01
UTAUT-predictors					
Social influence	.42	.06	0.30 to 0.54	.39	<.001
Performance expectancy	.31	.06	0.19 to 0.43	.31	<.001
Effort expectancy	.20	.06	.0.09 to 0.31	.22	<.001
Facilitating conditions	.06	.05	−0.03 to 0.16	.07	.19

[a] General self-efficacy short form (ASKU) [61].

[b] Patient Health Questionnaire-4 (PHQ-4) [65].

[c] Subjective prognosis of work ability scale (SPE) [66].

[d] eHealth literacy scale (eHEALS) [63].

Secondary Outcome

Overall, eHealth and IT literacy, knowledge of eHealth interventions, and self-efficacy were moderate. Mental distress ranged from “normal” (CAR) to “moderate” (PSY). Risk of incapacity to work was on average. Acceptance did not differ in extreme-group comparisons (PHQ-9: t_{285}=0.54, *P*=.59; SPE: t_{214}=1.48, *P*=.14). Participants reported low health-related Internet or mobile use, with less than 1 in 10 participants who indicated utilization for mental health problems or occupational stress. Significant differences of diagnostic groups were observed in all secondary outcome variables. Secondary outcome is reported in Table 4. Post-hoc tests revealed a distinguished role of the pediatric (higher IT literacy, eHealth literacy, eHealth knowledge, eHealth experience, lower ratings on stress through permanent availability, and Internet anxiety) and the psychosomatic subsample (low eHealth literacy, eHealth knowledge, eHealth experience, low self-efficacy, and high Internet anxiety). When including age as a covariate, these group differences remained significant.

Table 4. Means, standard deviations, and test statistics of group differences for secondary outcome measures (N=338).

Variable	PSY[a]	OPE[b]	CAR[c]	PED[d]	SUD[e]	Total	*F*	*P* value
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)		
Performance expectancy	2.41 (1.21)	2.26 (1.09)	2.59 (1.07)	3.24 (1.14)	2.34 (1.06)	2.54 (1.17)	$F_{4,333}=6.23$	P<.001
Effort expectancy	2.70 (1.29)	2.60 (1.13)	2.98 (1.16)	3.42 (1.03)	2.57 (1.20)	2.83 (1.22)	Welch $F_{4,134.92}=5.64$	P<.001
Social influence	2.52 (1.06)	2.41 (1.14)	2.79 (0.96)	2.87 (0.97)	2.38 (1.10)	2.59 (1.05)	$F_{4,333}=2.57$	.04
Facilitating conditions	3.09 (1.38)	2.95 (1.30)	3.30 (1.19)	3.98 (1.19)	2.91 (1.24)	3.22 (1.32)	$F_{4,333}=6.10$	P<.001
eHealth literacy[f]	3.30 (1.00)	3.63 (0.90)	3.76 (0.90)	4.08 (0.73)	3.37 (1.13)	3.56 (1.00)	Welch $F_{4,134.52}=9.37$	P<.001
Internet anxiety	2.57 (1.27)	2.28 (1.14)	2.10 (1.28)	1.87 (1.06)	2.24 (1.08)	2.28 (1.22)	$F_{4,333}=3.76$	P<.001
Self-efficacy[g]	3.10 (0.91)	4.07 (0.87)	4.11 (0.71)	3.50 (0.99)	3.67 (0.72)	3.57 (0.94)	Welch $F_{4,132.09}=20.75$	P<.001
Knowledge of eHealth interventions	2.43 (1.10)	3.08 (1.03)	2.90 (1.03)	3.28 (1.05)	2.73 (1.15)	2.78 (1.12)	$F_{4,333}=6.99$	P<.001
Health-related Internet and mobile use	1.38 (0.55)	1.51 (0.65)	1.71 (0.77)	1.97 (0.86)	1.55 (0.77)	1.58 (0.73)	Welch $F_{4,124.91}=6.28$	P<.001
Permanent availability stress	3.07 (1.37)	3.03 (1.26)	2.81 (1.31)	1.77 (1.08)	2.72 (1.37)	2.76 (1.37)	$F_{4,333}=9.55$	P<.001
Mental distress[h]	7.34 (3.21)	3.36 (3.24)	1.84 (2.29)	--	3.21 (2.85)	4.69 (3.77)	$F_{3,283}=61.72$	P<.001
Risk of incapacity to work[i]	1.41 (1.14)	1.38 (0.95)	0.89 (1.12)	--	1.07 (1.10)	1.22 (1.12)	$F_{3,278}=3.79$	.01
IT literacy	2.88 (1.19)	3.02 (1.14)	3.24 (1.18)	4.1 (0.76)	2.93 (1.25)	3.16 (1.2)	Welch $F_{4,134.68}=19.56$	P<.001

[a] PSY: psychosomatic medicine and psycho-oncology.

[b] OPE: orthopedics.

[c] CAR: cardiology.

[d] PED: pediatric disorders of adolescent patients.

[e] SUD: substance use disorders.

[f] eHealth literacy scale (eHEALS) [63].

[g] General self-efficacy short form (ASKU) [61]

[h] Patient Health Questionnaire-4 (PHQ-4) [65].

[i] Subjective prognosis of work ability scale (SPE) [66].

Furthermore, the majority of inpatients with valid answers (n=280) indicated interest in digital health-related information (68.9%, 193/280) and exercises (56.8%, 159/280). Half of the patients were interested in a Web-based expert contact (52.1%, 146/280) and the use of an app for aftercare (37.9%, 106/280). However, most participants with valid answers preferred face-to-face aftercare (53.4%, 179/335) over Web-based aftercare (4.8%, 16/335), 38.2% (128/335) rated Web-based aftercare as an add-on to traditional methods and 3.6% (12/335) had no interest in any form of aftercare (χ^2_3=247.99, P<.001).

Inter-rater reliability of open answers on advantages and disadvantages was high (Cohen kappa=.89, P<.001, 95% CI 0.54-1). Flexibility in terms of time and location of utilization were rated as main advantages, while impersonality and concerns about data security were perceived as main disadvantages of Web-based aftercare. Table 5 contains the most frequently reported advantages of 113 and disadvantages of 121 participants.

Table 5. Advantages and disadvantages of Web-based aftercare as measured by number of statements. Frequencies above 5% reported; infrequent statements aggregated in "others" category.

Category	n (%)
Advantages (n=152 statements)	
Flexibility in time	51 (33.6%)
Local flexibility	33 (21.7%)
Availability of personal support	11 (7.2%)
Reduced expenditure of time	9 (5.9%)
Availability and topicality of health information	8 (5.3%)
Anonymity	8 (5.3%)
Other	32 (21.1%)
Disadvantages (n=142 statements)	
Too impersonal	39 (27.5%)
Concerns about data security	14 (9.9%)
Increased expenditure of time	14 (9.9%)
Organizational conflicts	12 (8.5%)
Insufficient professional supervision	10 (7.0%)
Insufficient knowledge of Internet use	10 (7.0%)
General concerns about Internet use	8 (5.6%)
Insufficient motivation	8 (5.6%)
Other	27 (19.0%)

Discussion

Principal Findings

This study examined barriers and motivators to acceptance of Web-based aftercare in a transdiagnostic sample of inpatients building upon the UTAUT [23] and explored subgroup-specific effects.

The results indicated a rather low level of acceptance of Web-based aftercare for work-related stress. Taken together, 81.6% of the patients reported a low-to-moderate acceptance, which was comparable to previous findings in different clinical settings, such as inpatient diabetic [38], chronic pain [39], or primary care patients [40], and evidence from surveys in the general population [26,27]. Interestingly, acceptance ratings for Web-based aftercare were lower compared to uptake rates of face-to-face-aftercare [7,9]. This was confirmed by a clear face-to-face treatment preference of inpatients in this study, which corresponded to evidence from previous studies [26,27,72]. Despite growing evidence on the efficacy of eHealth interventions [1,2], acceptance in target groups seems limited, which points to different drivers and barriers to eHealth acceptance and adoption. However, high acceptance, eHealth literacy, and Internet-orientation in younger patients suggest sufficient resources to eHealth implementation in adolescent care. Since the incidence rates of mental disorders in adolescents are quite high [73], eHealth interventions thus seem promising in preventing or treating mental health problems.

This study supported the viability of the UTAUT model in assessing acceptance and found that social influence, performance, and effort expectancy, as well as stress through permanent availability, predicted acceptance. These factors may provide a framework for improving acceptance and implementation of eHealth interventions in inpatient rehabilitation.

The effect of *social influence* on acceptance was prominent, in contrast to some previous evidence, where this predictor reached smaller effects [48]. The influence of significant others in eHealth acceptance (family, friends, general practitioner) thus underlines the systemic aspect of the person-environment interaction in clinical settings. A positive attitude of significant others about eHealth efficacy and practitioners' willingness to refer these programs can foster adoption and adherence in patients. In particular, general practitioners may play an important role as gate keepers to eHealth adoption [74,75]. A study by van Voorhees et al [75] could demonstrate, for example, that uptake of a e-mental health intervention increased with client-centered information by clinicians, targeting intrinsic motivation. Therefore, facilitation of acceptance needs to include relevant stakeholders of health care (eg, practitioners, clinicians, administrators) as important mediators of eHealth implementation.

In line with previous research [47,48], we identified *performance expectancy* as a key predictor of acceptance, with negative outcome expectations predicting lower intention of utilization. Similarly, research suggests that performance expectancy may also be a predictor of treatment outcome in psychotherapy [76]. However, the disparity between low performance expectancy and actual efficacy of eHealth interventions [2] supports the need for further, transparent eHealth education, targeting common misconceptions, for example, about inferior efficacy in contrast to traditional therapies [26,27,72].

The influence of *effort expectancy* reflects the importance of a fit of usability and adoption in everyday life. This corroborates prior research, suggesting that incongruity of organization and scheduling of aftercare sessions with working life are key aspects of participation rate [7,11,35]. It seems that mechanisms of uptake are, to some extent, comparable between Web-based and face-to-face treatment. However, technical barriers and concerns about data security or impersonality as well as facilitators such as flexibility or anonymity are unique to eHealth interventions. Usability may relate to eHealth literacy as well [77], which was above average in this study and comparable to previous research in a student population [63]. However, future studies should expand the concept of eHealth literacy by investigating the ability to utilize eHealth interventions.

Interestingly, *stress through permanent availability* was associated with a lower acceptance, although the effect was very small. It would seem that Web-based aftercare may raise concerns about privacy or technological weariness in coping with health problems. In a qualitative study, Donkin and Glozier [78] identified technology fatigue as an important barrier to adherence to e-mental health interventions. A German survey of 538 employees found that 88% stated being reachable to clients or coworkers, and that 42% could not recognize a clear boundary between professional and private life [79]. Moreover, Reinecke et al [59] found that digital communication load was associated with burnout, depression, and anxiety in a representative survey. However, further studies are recommended to develop a full picture of the association between eHealth utilization, experience of stress through availability, and other sources of occupational stress.

Furthermore, our data show a differential influence of *demographics*, with a pattern of higher acceptance in younger, more educated, and eHealth literate patients with private Internet access and those with eHealth experience. This pattern confirms previous evidence on a relationship with Internet use [80], Internet competence [81], and acceptance of e-mental health interventions [27]. It may be that these mediate the influence of Internet affinity and prior use of eHealth interventions on acceptance. However, we could not find a continuous relation of acceptance and age. Nevertheless, our data point to a “digital divide” in eHealth adoption that may reflect existing social barriers to education or health promotion [82,83]. A recent survey, for example, revealed that only 30% of Germans older than 60 years used the Internet on a regular basis [80]. In this regard, Renahy et al [83] criticized limited access to eHealth not only in old aged but also in chronically ill patients. Future research should therefore aim to facilitate access and education and to adapt eHealth interventions to undersupplied or risk groups in particular.

However, other predictors did not prove to be relevant barriers or resources to eHealth acceptance in our sample. It would seem that a technical infrastructure is widely available and accessible to the majority of inpatients. As previously noted, attitudes may have a greater impact compared with structural barriers on utilization and adherence of eHealth interventions, in particular with ongoing technical progress [25]. However, eHealth intervention knowledge, eHealth experience, and health-related Internet and mobile use were low in most inpatients in our study. This can be applied to health professionals too, where previous studies found that only few practitioners offer e-mental health interventions to patients [84,85], thus limiting availability and exposure to effective treatment tools. Perle et al [52] found that 75% of clinical psychologists would reconsider using a Web camera–based e-mental health intervention after appropriate education. However, eHealth training for health professionals is scarce. When considering methods to facilitate acceptance, initial evidence in different target groups suggests that a brief video-based introduction or interactive presentation containing information about effectiveness, data security, or advantages of eHealth interventions can be an economic way to facilitate acceptance and resources to eHealth adoption [38-40,86]. Recently, Donovan et al [87] provided promising, yet mixed, evidence for a PowerPoint-based eHealth education in mental health professionals. However, further work is required to establish the viability of eHealth training and acceptance facilitation in different clinical contexts such as inpatient rehabilitation and to enhance familiarity and experience with eHealth interventions.

Limitations

Despite the strengths of our study, several limitations should be considered when interpreting the results. First, since we adapted items to the specific context of our study, psychometric validity cannot be fully ensured. To maintain parsimony, some predictors were operationalized with only 2 items, limiting their reliability. Also, differences in diagnostic groups may partly reflect diverging areas of application of Web-based aftercare between the adolescent and the adult populations. Future designs should increase item quantity to improve reliability of measurement and perform confirmatory factor or moderator analyses with larger sample sizes.

Second, the generalizability to other samples might be limited due to possible selection effects and diverging recruitment efficacy. This is reflected in a low overall response rate, ranging from 26% to 100% between facilities. In a way, the response rate may also represent an indicator of low acceptance of eHealth interventions. Furthermore, a large proportion of adult participants (36.4%) were not working before treatment, which might have limited acceptance of work-related interventions. Moreover, small sample sizes of some subgroup analyses (eg, diagnostic group comparisons, open answers) need to be considered. However, we were able to investigate a broad range of the most common diagnostic groups of German inpatient rehabilitation [88] without preselection of Internet-oriented individuals. Our study relied on a well-established theoretical model and validated secondary outcome measures and proved their reliability in, and applicability to, a mixed inpatient sample. However, future research should promote reliable instruments of acceptance evaluation in different contexts. The aforementioned APOI [37] found the factors “Scepticism and Perception of Risks,” “Confidence in Effectiveness” (similar to performance expectancy), “Technologization Threat” (similar to stress through constant availability), and “Anonymity Benefits” (p. 140) as dimensions of attitudes toward e-mental-health interventions. Moreover, Boß et al [89] recently adapted the Client Satisfaction Questionnaire [90] to eHealth interventions. Also, Mohr et al [91] developed a 24-item questionnaire measuring perceived barriers to psychological treatment, which could be adapted to the context of eHealth interventions. Prospectively, these could also provide clinicians with reliable and valuable information in the implementation of eHealth interventions in practice.

Third, it should be noted that in line with previous studies, we measured acceptance as behavioral intention, which allows only for a prediction but not for direct translation into actual use behavior, given the well-known intention-behavior gap [92]. Future research should therefore include uptake rate as an outcome measure to improve interpretability of motivational aspects and allow for a direct comparison of utilization of Web-based and face-to-face aftercare. Fourth, not only motivational but also structural barriers to eHealth adoption need to be considered when investigating acceptance. Institutional or technical constraints may limit application of eHealth interventions [93]. In Germany and other countries, the use of eHealth interventions still underlies legal constraints, requiring at least an initial face-to-face contact with practitioners. Thus, even if willingness to use eHealth interventions was higher, uptake could be limited by such macro-level constraints.

Finally, the cross-sectional design of the study does not allow for a direct projection to further technical developments and progress of implementation of eHealth interventions. Longitudinal designs will be needed to accompany technical progress and change in underlying barriers and facilitators to acceptance.

Conclusions

Our study contributes to the exploration of barriers and drivers to eHealth implementation in inpatient treatment and illustrates a limited but developable acceptance. Our results suggest that expectations on efficacy and usability, social norms, and experience of permanent availability modulate acceptance of Web-based aftercare. Thus, it seems critical not only to increase eHealth experience and literacy but to facilitate positive attitudes and target misconceptions about eHealth interventions, regarding competitiveness to face-to-face treatment, effectiveness, individuality, or therapeutic relationship. More elaborate forms of education and quality assurance in implementation are needed, for example, with the development of mandatory and comprehensible quality criteria or training in the application of eHealth interventions for health professionals. Future developments of eHealth interventions therefore should include patients as well as other stakeholders in a collaborative approach to allow for a better conjunction between technical, contextual, and motivational factors influencing adoption and effectiveness of eHealth interventions.

Acknowledgments

Our study was part of a dissertation project of SH and funded by a research grant of the Illa and Werner Zarnekow Foundation for rehabilitation sciences. The funders had no role in designing the study, data collection and analysis, or preparation of this manuscript. We would like to thank all participants of the survey and the clinical directors and research assistants of the Mittelrhein-Klinik Boppard/Bad Salzig (Frank M Rudolph, MD; Ursula Wittig), Edelsteinklinik Bruchweiler (Edith Waldeck, MD; Daniel Funck), Drei-Burgen-Klinik Bad Kreuznach (Ute Dederichs-Masius, MD; Michael Wiegert), and Fachklinik Eußerthal (Thomas Korte, MD; Nikolaus Lange).

Authors' Contributions

RZ initiated the study. SH designed and conducted the study and wrote the draft of the manuscript. RZ and MB contributed to the design of this study. All authors contributed to the further writing of the manuscript and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Exemplary description of different Web-based aftercare interventions following inpatient treatment.

[XLSX File (Microsoft Excel File), 14KB - jmir_v18i12e337_app1.xlsx]

Multimedia Appendix 2

Characteristics of the sample.

[XLSX File (Microsoft Excel File), 17KB - jmir_v18i12e337_app2.xlsx]

Multimedia Appendix 3

Correlation matrix of acceptance predictors.

[XLSX File (Microsoft Excel File), 15KB - jmir_v18i12e337_app3.xlsx]

References

1. 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]
2. Barak A, Hen L, Boniel-Nissim M, Shapira N. A comprehensive review and a meta-analysis of the effectiveness of Internet-based psychotherapeutic interventions. J Technol Hum Serv 2008 Jul 03;26(2-4):109-160. [doi: 10.1080/15228830802094429]
3. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
4. Hedman E, Andersson E, Ljótsson B, Andersson G, Rück C, Lindefors N. Cost-effectiveness of Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: results from a randomized controlled trial. Behav Res Ther 2011 Nov;49(11):729-736. [doi: 10.1016/j.brat.2011.07.009] [Medline: 21851929]
5. Lin J, Ebert DD, Lehr D, Berking M, Baumeister H. [Internet based cognitive behavioral interventions: state of the art and implementation possibilities in rehabilitation]. Rehabilitation 2013 Jun;52(3):155-163. [doi: 10.1055/s-0033-1343491] [Medline: 23761204]
6. Gönner S, Bischoff C, Ehrhardt M, Limbacher K. [Effects of therapy goal oriented cognitive behavioural aftercare measures on therapy transfer following inpatient psychosomatic rehabilitation]. Rehabilitation 2006 Dec;45(6):369-376. [doi: 10.1055/s-2006-932614] [Medline: 17123219]
7. Sibold M, Mittag O, Kulick B, Müller E, Opitz U, Jäckel WH. [Predictors of participation in medical rehabilitation follow-up in working patients with chronic back pain]. Rehabilitation 2011 Dec;50(6):363-371. [doi: 10.1055/s-0031-1271815] [Medline: 21647850]
8. Kobelt A, Nickel L, Grosch EV, Lamprecht F, Künsebeck H. [Participation in psychosomatic outpatient care after in-patient rehabilitation]. Psychother Psychosom Med Psychol 2004 Feb;54(2):58-64. [doi: 10.1055/s-2003-812612] [Medline: 14872392]
9. Köpke K. [Upgrade, expand, systematize--an analysis of the status of, the need for reform in, as well as innovative projects for follow-up care in rehabilitation under the German pension insurance scheme]. Rehabilitation 2005 Dec;44(6):344-352. [doi: 10.1055/s-2005-915247] [Medline: 16320178]
10. Ebert D, Tarnowski T, Gollwitzer M, Sieland B, Berking M. A transdiagnostic internet-based maintenance treatment enhances the stability of outcome after inpatient cognitive behavioral therapy: a randomized controlled trial. Psychother Psychosom 2013;82(4):246-256. [doi: 10.1159/000345967] [Medline: 23736751]
11. Ebert DD, Hannig W, Tarnowski T, Sieland B, Götzky B, Berking M. [Web-based rehabilitation aftercare following inpatient psychosomatic treatment]. Rehabilitation 2013 Jun;52(3):164-172. [doi: 10.1055/s-0033-1345191] [Medline: 23761205]
12. Schmädeke S, Bischoff C. [Effects of smartphone-supported rehabilitation aftercare (eATROS) for depressive patients]. Verhaltenstherapie 2015;25(4):277-286. [doi: 10.1159/000441856]
13. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions - A systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
14. Moessner M, Aufdermauer N, Baier C, Göbel H, Kuhnt O, Neubauer E, et al. [Efficacy of an internet-delivered aftercare program for patients with chronic back pain]. Psychother Psychosom Med Psychol 2014 Feb;64(2):47-53. [doi: 10.1055/s-0033-1351266] [Medline: 24101036]
15. Zwerenz R, Gerzymisch K, Edinger J, Holme M, Knickenberg RJ, Spörl-Dönch S, et al. Evaluation of an internet-based aftercare program to improve vocational reintegration after inpatient medical rehabilitation: study protocol for a cluster-randomized controlled trial. Trials 2013 Jan 25;14:26 [FREE Full text] [doi: 10.1186/1745-6215-14-26] [Medline: 23351836]
16. Fledderus M, Schreurs KM, Bohlmeijer ET, Vollenbroek-Hutten MM. Development and pilot evaluation of an online relapse-prevention program based on acceptance and commitment therapy for chronic pain patients. JMIR Hum Factors 2015 Jan 05;2(1):e1 [FREE Full text] [doi: 10.2196/humanfactors.3302] [Medline: 27025550]
17. Zwerenz R, Knickenberg RJ, Schattenburg L, Beutel ME. [Work-related stress and resources of psychosomatic patients compared to the general population]. Rehabilitation 2004 Feb;43(1):10-16. [doi: 10.1055/s-2004-818547] [Medline: 14963800]
18. Geraedts AS, Kleiboer AM, Twisk J, Wiezer NM, van Mechelen W, Cuijpers P. Long-term results of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014;16(7):e168 [FREE Full text] [doi: 10.2196/jmir.3539] [Medline: 25008127]

19. Ebert DD, Lehr D, Smit F, Zarski A, Riper H, Heber E, et al. Efficacy and cost-effectiveness of minimal guided and unguided internet-based mobile supported stress-management in employees with occupational stress: a three-armed randomised controlled trial. BMC Public Health 2014;14:807 [FREE Full text] [doi: 10.1186/1471-2458-14-807] [Medline: 25099533]
20. Eichenberg C, Ott R. [Internet-based psychological intervention. Review of empirical results of disorder-specific programs]. Psychotherapeut 2011;57(1):58-69. [doi: 10.1007/s00278-011-0832-5]
21. Kaltenthaler E, Sutcliffe P, Parry G, Beverley C, Rees A, Ferriter M. The acceptability to patients of computerized cognitive behaviour therapy for depression: a systematic review. Psychol Med 2008 Nov;38(11):1521-1530. [doi: 10.1017/S0033291707002607] [Medline: 18205964]
22. 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]
23. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478. [doi: 10.2307/30036540]
24. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319. [doi: 10.2307/249008]
25. Waller R, Gilbody S. Barriers to the uptake of computerized cognitive behavioural therapy: a systematic review of the quantitative and qualitative evidence. Psychol Med 2009 May;39(5):705-712. [doi: 10.1017/S0033291708004224] [Medline: 18812006]
26. Gun SY, Titov N, Andrews G. Acceptability of Internet treatment of anxiety and depression. Australas Psychiatry 2011 Jun;19(3):259-264 [FREE Full text] [doi: 10.3109/10398562.2011.562295] [Medline: 21682626]
27. Eichenberg C, Wolters C, Brähler E. The Internet as a Mental Health Advisor in Germany—Results of a National Survey. PLoS One 2013;8(11):e79206 [FREE Full text] [doi: 10.1371/journal.pone.0079206] [Medline: 24278121]
28. Kobelt A, Karpinski N, Petermann F. [Internet-based assessment in medical rehabilitation: patient's acceptance]. Phys Rehab Kur Med 2010;20(06):316-321. [doi: 10.1055/s-0030-1267243]
29. Tavares J, Oliveira T. Electronic health record patient portal adoption by health care consumers: an acceptance model and survey. J Med Internet Res 2016 Mar 02;18(3):e49 [FREE Full text] [doi: 10.2196/jmir.5069] [Medline: 26935646]
30. Bartel S, Nowossadeck E, Spyra K. [Acceptance of internet-based aftercare elements in orthopedic and cardiological inpatients]. In: DRV-Schriften, Vol. 88. Berlin: Deutsche Rentenversicherung Bund; 2010 Presented at: 19th Rehabilitationswissenschaftliches Kolloquium; 2010, March 8-10; Leipzig, Germany p. 175-176 URL: http://forschung.deutsche-rentenversicherung.de/ForschPortalWeb/ressource?key=tagungsband_19_reha_kolloqu.pdf
31. Wild B, Hünnemeyer K, Sauer H, Hain B, Mack I, Schellberg D, et al. A 1-year videoconferencing-based psychoeducational group intervention following bariatric surgery: results of a randomized controlled study. Surg Obes Relat Dis 2015;11(6):1349-1360. [doi: 10.1016/j.soard.2015.05.018] [Medline: 26421929]
32. Kramer D, Moessner M, Görlitz S, Bauer S. [Mobile and internet-based aftercare for alcohol addicts following inpatient withdrawal therapy]. In: DRV-Schriften, Vol. 109. Berlin: Deutsche Rentenversicherung Bund; 2016 Presented at: 25th Rehabilitationswissenschaftliches Kolloquium; 2016 Feb 29 - Mar 02; Aachen, Germany p. 147-148 URL: http://forschung.deutsche-rentenversicherung.de/ForschPortalWeb/ressource?key=Band_109_Internet.pdf
33. Pfaudler S, Hoberg E, Benninghoven D. [Internet-based aftercare for patients after multimodal orthopedic rehabilitation - feasibility and acceptance]. Phys Rehab Kur Med 2015;25(02):81-89. [doi: 10.1055/s-0034-1396817]
34. Theissing J, Deck R, Raspe H. [Liveonline aftercare in patients with abdominal obesity in cardio-diabetological rehabilitation: findings of a randomized controlled study]. Rehabilitation 2013 Jun;52(3):153-154. [doi: 10.1055/s-0033-1345190] [Medline: 23761202]
35. Küffner R. [Live-online-aftercare]. 2010. URL: http://www.psychotherapie.uni-wuerzburg.de/forschung/Abschlussbericht_Live-Online-Nachsorge.pdf [accessed 2016-10-07] [WebCite Cache ID 6l5H0B7HX]
36. Fichter MM, Quadflieg N, Lindner S. Internet-based relapse prevention for anorexia nervosa: nine- month follow-up. J Eat Disord 2013;1:23 [FREE Full text] [doi: 10.1186/2050-2974-1-23] [Medline: 24999404]
37. Schröder J, Sautier L, Kriston L, Berger T, Meyer B, Späth C, et al. Development of a questionnaire measuring attitudes towards psychological online interventions-the APOI. J Affect Disord 2015 Nov 15;187:136-141. [doi: 10.1016/j.jad.2015.08.044] [Medline: 26331687]
38. Baumeister H, Nowoczin L, Lin J, Seifferth H, Seufert J, Laubner K, et al. Impact of an acceptance facilitating intervention on diabetes patients' acceptance of Internet-based interventions for depression: a randomized controlled trial. Diabetes Res Clin Pract 2014 Jul;105(1):30-39. [doi: 10.1016/j.diabres.2014.04.031] [Medline: 24862240]
39. Baumeister H, Seifferth H, Lin J, Nowoczin L, Lüking M, Ebert D. Impact of an Acceptance Facilitating Intervention on Patients' Acceptance of Internet-based Pain Interventions: A Randomized Controlled Trial. Clin J Pain 2015 Jun;31(6):528-535. [doi: 10.1097/AJP.0000000000000118] [Medline: 24866854]
40. Ebert DD, Berking M, Cuijpers P, Lehr D, Pörtner M, Baumeister H. Increasing the acceptance of internet-based mental health interventions in primary care patients with depressive symptoms. A randomized controlled trial. J Affect Disord 2015 May 1;176:9-17. [doi: 10.1016/j.jad.2015.01.056] [Medline: 25682378]
41. Wilson EV, Lankton NK. Modeling patients' acceptance of provider-delivered e-health. J Am Med Inform Assoc 2004;11(4):241-248 [FREE Full text] [doi: 10.1197/jamia.M1475] [Medline: 15064290]

42. Li J, Talaei-Khoei A, Seale H, Ray P, Macintyre CR. Health care provider adoption of eHealth: systematic literature review. Interact J Med Res 2013;2(1):e7 [FREE Full text] [doi: 10.2196/ijmr.2468] [Medline: 23608679]
43. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
44. Dünnebeil S, Sunyaev A, Blohm I, Leimeister JM, Krcmar H. Determinants of physicians' technology acceptance for e-health in ambulatory care. Int J Med Inform 2012 Nov;81(11):746-760. [doi: 10.1016/j.ijmedinf.2012.02.002] [Medline: 22397989]
45. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
46. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
47. Liu L, Miguel CA, Rios RA, Buttar V, Ranson Q, Goertzen D. What factors determine therapists' acceptance of new technologies for rehabilitation - a study using the Unified Theory of Acceptance and Use of Technology (UTAUT). Disabil Rehabil 2015;37(5):447-455. [doi: 10.3109/09638288.2014.923529] [Medline: 24901351]
48. Taiwo A, Downe A. The theory of user acceptance and use of technology (UTAUT): a meta-analytic review of empirical findings. Journal of Theoretical & Applied Information Technology 2013;49(1):48-58.
49. Gerhards SA, Abma TA, Arntz A, de Graaf LE, Evers SM, Huibers MJ, et al. Improving adherence and effectiveness of computerised cognitive behavioural therapy without support for depression: a qualitative study on patient experiences. J Affect Disord 2011 Mar;129(1-3):117-125. [doi: 10.1016/j.jad.2010.09.012] [Medline: 20889214]
50. Kontos E, Blake KD, Chou WY, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
51. Celik V, Yesilyurt E. Attitudes to technology, perceived computer self-efficacy and computer anxiety as predictors of computer supported education. Comput Educ 2013 Jan;60(1):148-158. [doi: 10.1016/j.compedu.2012.06.008]
52. Perle JG, Langsam LC, Randel A, Lutchman S, Levine AB, Odland AP, et al. Attitudes toward psychological telehealth: current and future clinical psychologists' opinions of internet-based interventions. J Clin Psychol 2013 Jan;69(1):100-113. [doi: 10.1002/jclp.21912] [Medline: 22975897]
53. Currie M, Philip LJ, Roberts A. Attitudes towards the use and acceptance of eHealth technologies: a case study of older adults living with chronic pain and implications for rural healthcare. BMC Health Serv Res 2015 Apr 16;15:162 [FREE Full text] [doi: 10.1186/s12913-015-0825-0] [Medline: 25888988]
54. Cranen K, Drossaert CH, Brinkman ES, Braakman-Jansen AL, Ijzerman MJ, Vollenbroek-Hutten MM. An exploration of chronic pain patients' perceptions of home telerehabilitation services. Health Expect 2012 Dec;15(4):339-350. [doi: 10.1111/j.1369-7625.2011.00668.x] [Medline: 21348905]
55. Young KS. An empirical examination of client attitudes towards online counseling. Cyberpsychol Behav 2005 Apr;8(2):172-177. [doi: 10.1089/cpb.2005.8.172] [Medline: 15938657]
56. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
57. Clarke J, Proudfoot J, Birch M, Whitton AE, Parker G, Manicavasagar V, et al. Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: secondary analysis of a randomised controlled trial. BMC Psychiatry 2014;14:272 [FREE Full text] [doi: 10.1186/s12888-014-0272-1] [Medline: 25252853]
58. Christensen H, Griffiths KM, Korten AE, Brittliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. J Med Internet Res 2004 Dec 22;6(4):e46 [FREE Full text] [doi: 10.2196/jmir.6.4.e46] [Medline: 15631970]
59. Reinecke L, Aufenanger S, Beutel ME, Dreier M, Quiring O, Stark B, et al. Digital stress over the life span: the effects of communication load and internet multitasking on perceived stress and psychological health impairments in a German probability sample. Media Psychol 2016 Mar:1-26. [doi: 10.1080/15213269.2015.1121832]
60. Eisinga R, Grotenhuis MT, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? Int J Public Health 2013 Aug;58(4):637-642. [doi: 10.1007/s00038-012-0416-3] [Medline: 23089674]
61. Beierlein C, Kemper C, Kovaleva A, Rammstedt B. [Short Scale for Measuring General Self-efficacy Beliefs (ASKU)]. Methoden, Daten, Analysen 2013;7(2):251-278 [FREE Full text] [doi: 10.12758/mda.2013.014]
62. Blankers M, Koeter MW, 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]
63. Soellner R, Huber S, Reder M. The Concept of eHealth Literacy and Its Measurement. J Media Psychol 2014 Jan;26(1):29-38. [doi: 10.1027/1864-1105/a000104]
64. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

65. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621 [FREE Full text] [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
66. Mittag O, Raspe H. [A brief scale for measuring subjective prognosis of gainful employment: findings of a study of 4279 statutory pension insurees concerning reliability (Guttman scaling) and validity of the scale]. Rehabilitation 2003 Jun;42(3):169-174. [doi: 10.1055/s-2003-40095] [Medline: 12813654]
67. Mittag O, Glaser-Möller N, Ekkernkamp M, Matthis C, Héon-Klin V, Raspe A, et al. [Predictive validity of a brief scale to assess subjective prognosis of work capacity (SPE Scale) in a cohort of LVA insured patients with severe back pain or functional complaints relating to internal medicine]. Sozial- und Präventivmedizin 2003 Dec 1;48(6):361-369. [doi: 10.1007/s00038-003-2037-3]
68. Tuomi K, Ilmarinen J, Jahkola A, Katajarinne L, Tulkki A. Arbeitsbewältigungsindex - Work Ability Index. Bremerhaven: Wirtschaftsverlag NW Verlag für Neue Wissenschaft; 2001.
69. Enders CK. Applied Missing Data Analysis. New York: The Guilford Press; 2010.
70. Midi H, Sarkar S, Rana S. Collinearity diagnostics of binary logistic regression model. Journal of Interdisciplinary Mathematics 2010 Jun;13(3):253-267. [doi: 10.1080/09720502.2010.10700699]
71. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. Weinheim: Beltz; 2015.
72. 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]
73. Bor W, Dean AJ, Najman J, Hayatbakhsh R. Are child and adolescent mental health problems increasing in the 21st century? A systematic review. Aust N Z J Psychiatry 2014 Jul;48(7):606-616 [FREE Full text] [doi: 10.1177/0004867414533834] [Medline: 24829198]
74. Lillevoll KR, Wilhelmsen M, Kolstrup N, Høifødt RS, Waterloo K, Eisemann M, et al. Patients' experiences of helpfulness in guided internet-based treatment for depression: qualitative study of integrated therapeutic dimensions. J Med Internet Res 2013;15(6):e126 [FREE Full text] [doi: 10.2196/jmir.2531] [Medline: 23786763]
75. Van Voorhees BW, Hsiung RC, Marko-Holguin M, Houston TK, Fogel J, Lee R, et al. Internal versus external motivation in referral of primary care patients with depression to an internet support group: randomized controlled trial. J Med Internet Res 2013 Mar 12;15(3):e42 [FREE Full text] [doi: 10.2196/jmir.2197] [Medline: 23482332]
76. Constantino MJ, Arnkoff DB, Glass CR, Ametrano RM, Smith JZ. Expectations. J Clin Psychol 2011 Feb;67(2):184-192. [doi: 10.1002/jclp.20754] [Medline: 21128304]
77. Stellefson M, Hanik B, Chaney B, Chaney D, Tennant B, Chavarria EA. eHealth literacy among college students: a systematic review with implications for eHealth education. J Med Internet Res 2011 Dec;13(4):e102 [FREE Full text] [doi: 10.2196/jmir.1703] [Medline: 22155629]
78. Donkin L, Glozier N. Motivators and motivations to persist with online psychological interventions: a qualitative study of treatment completers. J Med Internet Res 2012;14(3):e91 [FREE Full text] [doi: 10.2196/jmir.2100] [Medline: 22743581]
79. Netzgesellschaft Eine repräsentative Untersuchung zur Mediennutzung und dem Informationsverhalten der Gesellschaft in Deutschland. 2011. URL: https://www.bitkom.org/Publikationen/2011/Studie/Studie-Netzgesellschaft/BITKOM-Publikation-Netzgesellschaft.pdf [accessed 2016-10-24] [WebCite Cache ID 6lV4MruaX]
80. Frees B, Koch W. Internetnutzung: Frequenz und Vielfalt nehmen in allen Altersgruppen zu. Media Perspektiven. 2015. URL: http://www.ard-zdf-onlinestudie.de/fileadmin/Onlinestudie_2015/0915_Frees_Koch.pdf [accessed 2016-12-22] [WebCite Cache ID 6mwTBuSWZ]
81. van Deursen AJ, van Dijk JA, Peters O. Rethinking internet skills: the contribution of gender, age, education, internet experience, and hours online to medium- and content-related Internet skills. Poetics 2011 Apr;39(2):125-144. [doi: 10.1016/j.poetic.2011.02.001]
82. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
83. Renahy E, Parizot I, Chauvin P. Health information seeking on the Internet: a double divide? Results from a representative survey in the Paris metropolitan area, France, 2005-2006. BMC Public Health 2008;8:69 [FREE Full text] [doi: 10.1186/1471-2458-8-69] [Medline: 18291024]
84. MacLeod M, Martinez R, Williams C. Cognitive Behaviour Therapy self-help: who does it help and what are its drawbacks? Behav Cogn Psychother 2009 Jan;37(1):61-72. [doi: 10.1017/S1352465808005031] [Medline: 19364408]
85. Eichenberg C, Kienzle K. [Psychotherapists and the Internet. Attitude toward therapeutic online offers and their use in routine treatment]. Psychotherapeut 2013 Sep;58(5):485-493. [doi: 10.1007/s00278-011-0833-4]
86. Mitchell N, Gordon PK. Attitudes towards computerized CBT for depression amongst a student population. Behav Cognit Psychother 2007 May 14;35(04):421. [doi: 10.1017/s1352465807003700]
87. Donovan CL, Poole C, Boyes N, Redgate J, March S. Australian mental health worker attitudes towards cCBT: What is the role of knowledge? Are there differences? Can we change them? Internet Interv 2015 Nov;2(4):372-381. [doi: 10.1016/j.invent.2015.09.001]
88. Reha-Bericht: die medizinische und berufliche Rehabilitation der Rentenversicherung im Licht der Statistik. Berlin: Deutsche Rentenversicherung Bund; 2015. URL: http://www.deutsche-rentenversicherung.de/Allgemein/de/Inhalt/6_Wir_ueber_uns/

03_fakten_und_zahlen/03_statistiken/02_statistikpublikationen/02_rehabericht_2015.pdf?__blob=publicationFile&v=4 [accessed 2016-12-19] [WebCite Cache ID 6l4vznO5n]
89. Boß L, Lehr D, Reis D, Vis C, Riper H, Berking M, et al. Reliability and validity of assessing user satisfaction with web-based health interventions. J Med Internet Res 2016 Aug 31;18(8):e234 [FREE Full text] [doi: 10.2196/jmir.5952] [Medline: 27582341]
90. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979;2(3):197-207. [Medline: 10245370]
91. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
92. Bhattacherjee A, Sanford C. The intention–behaviour gap in technology usage: the moderating role of attitude strength. Behav Inf Technol 2009 Jul;28(4):389-401. [doi: 10.1080/01449290802121230]
93. Eichenberg C, Molitor K. [Application of different media in inpatient psychotherapy. Results of a study surveying therapists and patients]. Psychotherapeut 2011 Mar;56(2):162-170. [doi: 10.1007/s00278-010-0736-9]

Abbreviations

ASKU: general self-efficacy short form
CAR: cardiology
CBT: cognitive behavioral therapy
eHEALS: eHealth literacy scale
OPE: orthopedics
PED: pediatric disorders of adolescent patients
PSY: psychosomatic medicine and psycho-oncology
PHQ-4: Patient Health Questionnaire-4
SPE: subjective prognosis of work ability scale
SUD: substance use disorders
UTAUT: unified theory of acceptance and use of technology

Edited by G Eysenbach; submitted 20.05.16; peer-reviewed by C Eichenberg, E Neter, J Tavares, M Moessner; comments to author 24.08.16; revised version received 07.10.16; accepted 22.10.16; published 23.12.16.

Please cite as:
Hennemann S, Beutel ME, Zwerenz R
Drivers and Barriers to Acceptance of Web-Based Aftercare of Patients in Inpatient Routine Care: A Cross-Sectional Survey
J Med Internet Res 2016;18(12):e337
URL: http://www.jmir.org/2016/12/e337/
doi:10.2196/jmir.6003
PMID:28011445

Original Paper

Influence of Pokémon Go on Physical Activity: Study and Implications

Tim Althoff[1], MS; Ryen W White[2], PhD; Eric Horvitz[2], MD, PhD

[1]Computer Science Department, Stanford University, Stanford, CA, United States

[2]Microsoft Research, Redmond, WA, United States

Corresponding Author:
Tim Althoff, MS
Computer Science Department
Stanford University
94 Thoburn Ct.
Apt 108
Stanford, CA, 94305
United States
Phone: 1 6504850758
Fax: 1 650 725 7411
Email: althoff@cs.stanford.edu

Abstract

Background: Physical activity helps people maintain a healthy weight and reduces the risk for several chronic diseases. Although this knowledge is widely recognized, adults and children in many countries around the world do not get recommended amounts of physical activity. Although many interventions are found to be ineffective at increasing physical activity or reaching inactive populations, there have been anecdotal reports of increased physical activity due to novel mobile games that embed game play in the physical world. The most recent and salient example of such a game is Pokémon Go, which has reportedly reached tens of millions of users in the United States and worldwide.

Objective: The objective of this study was to quantify the impact of Pokémon Go on physical activity.

Methods: We study the effect of Pokémon Go on physical activity through a combination of signals from large-scale corpora of wearable sensor data and search engine logs for 32,000 Microsoft Band users over a period of 3 months. Pokémon Go players are identified through search engine queries and physical activity is measured through accelerometers.

Results: We find that Pokémon Go leads to significant increases in physical activity over a period of 30 days, with particularly engaged users (ie, those making multiple search queries for details about game usage) increasing their activity by 1473 steps a day on average, a more than 25% increase compared with their prior activity level ($P<.001$). In the short time span of the study, we estimate that Pokémon Go has added a total of 144 billion steps to US physical activity. Furthermore, Pokémon Go has been able to increase physical activity across men and women of all ages, weight status, and prior activity levels showing this form of game leads to increases in physical activity with significant implications for public health. In particular, we find that Pokémon Go is able to reach low activity populations, whereas all 4 leading mobile health apps studied in this work largely draw from an already very active population.

Conclusions: Mobile apps combining game play with physical activity lead to substantial short-term activity increases and, in contrast to many existing interventions and mobile health apps, have the potential to reach activity-poor populations. Future studies are needed to investigate potential long-term effects of these applications.

(J Med Internet Res 2016;18(12):e315) doi:10.2196/jmir.6759

KEYWORDS

physical activity; Pokémon Go; mobile health; mHealth; wearable devices; mobile applications; games; exergames; public health

Introduction

“Those who think they have no time for bodily exercise will sooner or later have to find time for illness.” [Edward Stanley, Earl of Derby, December 20, 1873]

Physical activity is critical to human health. People who are physically active tend to live longer, have lower risk for heart disease, stroke, type 2 diabetes, depression, and some cancers, and are more likely to maintain a healthy weight [1,2,3,4]. Recent analyses estimate that physical inactivity contributes to 5.3 million deaths per year worldwide [5] and that it is responsible for a worldwide economic burden of US $67.5 billion due to health care expenditures and productivity losses [6]. Only 21% of US adults meet official physical activity guidelines [7,8] (at least 150 minutes a week of physical activity for adults), and less than 30% of US high school students get at least 60 minutes of physical activity every day [7]. Efforts to stimulate physical activity hold opportunity for improving public health. Numerous studies have called for population-wide approaches [9,10]. However, many interventions have been found to be either ineffective [11,12] in reaching only populations that were already active [13,14] or not scalable across varying cultural, geographic, social, and economic contexts [9].

Of late, there have been anecdotal reports of novel mobile games leading to increased physical activity, most notably for Pokémon Go [15] (other examples include Ingress and Zombies, Run!). Pokémon Go is a mobile game combining the Pokémon world through augmented reality with the real world requiring players to physically move around. Pokémon Go was released in the United States on July 6, 2016 and was adopted widely around the world—25 million active users in the United States [16] and 40 million worldwide [17]; 500 million downloads worldwide [18]. Due to this massive penetration, Pokémon Go can be viewed as an intervention for physical activity on a large scale. However, its effectiveness for stimulating additional walking is yet to be determined.

Our study seeks to provide guidance on large-scale interventions represented by the Pokémon Go phenomenon and on the possibilities for increasing physical activity which could be leveraged games for public health purposes. We study the influence of Pokémon Go on physical activity through a combination of wearable sensor data and search engine query logs for 31,793 users over a period of 3 months. Within these users, we identify 1420 Pokémon Go users based on their search activity and measure the effect of playing the game on their physical activity. We compare changes in physical activity for Pokémon Go users with changes for a large control group of US wearable users and further compare the effect of Pokémon Go with 4 leading mobile health apps. Finally, we estimate the impact of Pokémon Go on public health.

To summarize, our main research questions are

1. Is playing Pokémon Go associated with increases in physical activity? How large is this effect and how long does it persist?

2. Is this effect restricted to particular subpopulations or is it affecting people of all prior activity levels, ages, gender, and weight status?

3. How does Pokémon Go compare with leading mobile health apps in terms of its ability to change physical activity?

4. How has Pokémon Go affected physical activity in the United States and what is its potential impact on public if the game was able to sustain the engagement of its users?

To the best of our knowledge, this is the first study to combine large-scale wearable and search sensors to retrospectively evaluate large-scale interventions and the first to study the effect of Pokémon Go.

Methods

Dataset

We combine data from search engine queries with physical activity measurements from wearable devices. Specifically, we jointly analyze (1) queries to the Microsoft Bing search engine mentioning “pokemon.” We use this to identify which users are likely playing Pokémon Go; and (2) physical activity as measured through daily number of steps on the Microsoft Band. We jointly use these data to measure differences in physical activity before and after each user shows strong evidence of starting to play Pokémon Go.

The main study population is the 31,793 US users of Microsoft products who have agreed to link data from their Microsoft Band wearables and their online activities to understand product usage and improve Microsoft products. We show that 1420 users can be classified as Pokémon Go players with high confidence. We compare changes in physical activity in this population with changes in a control group consisting of a random sample of 50,000 US Microsoft Band users. As the expected fraction of Pokémon Go players in this random sample is very low [19], we are confident that this does not affect our analyses. For all users, we have self-reported age, gender, height, and weight, which are used to estimate the effect of Pokémon Go on different groups of users.

This research was conducted while all authors were employed by Microsoft. Our work was conducted offline, on data collected to support existing business operations, and did not influence the user experience. All data were anonymized prior to analyses. The Ethics Advisory Committee at Microsoft Research considers these precautions sufficient for triggering the Common Rule, exempting this work from detailed ethics review.

Table 1. Representative experiential and nonexperiential Pokémon Go queries [20]. In total, 454 such features were used for classification of experiential queries.

Nonexperiential query	Experiential query
pokemon go	pokemon go iv[a] calculator
pokemon go death san francisco	pokemon go teams
pokemon go robberies	how to play pokemon go
couple sues pokemon go	pokemon go guide
baltimore pokemon accident	pokemon go servers
pokemon games	pokemon go bot
bluestacks[b] pokemon go	pokemon go eevee[c] evolution

[a] iv: It refers to individual values which are attribute points of Pokémon determining their stamina, attach, and strength.

[b] bluestacks: It refers to a method to play Pokémon Go on a desktop computer instead of the intended use in the real world.

[c] eevee: It is the name of a Pokémon.

Identification of Pokémon Go Users Through Search Queries

We collected all Bing search queries from any device (eg, desktop, laptop, or mobile) of the 31,793 users between July 6, 2016 (US release date of Pokémon Go) and August 23, 2016 (date of statistical analysis) that mention the term “pokemon” (ignoring capitalization). We then manually annotated the 454 most frequent unique queries in terms of whether they are experiential [20,21]; that is, the user is very likely playing Pokémon Go, rather than just being interested in it for some other reason such as following up on news reports or general interest in the game. This was done by an author familiar with the game manually executing each query and judging whether the query and search engine results provided compelling evidence of someone playing the game. Examples for experiential and non-experiential queries are given in Table 1.

Among the 25,446 users who issued any queries during our time of observation, 1420 or 5.58% issued an experiential query for Pokémon Go. This number closely matches the estimated fraction of regular Pokémon users in the United States—5.9% according to [19]—suggesting that our search-engine based method is effectively detecting a large number of Pokémon Go users. We use the time of each user’s first experiential query for Pokémon Go as a proxy for the time when they started playing Pokémon Go and denote this time as t_0.

Note that our method of identifying Pokémon Go players through experiential queries can potentially overestimate t_0, if players perform these queries several days after starting to play the game, but the opposite is less likely due to the nature of experiential queries targeting specific aspects of game play (see Table 1). However, note that any potential overestimates of t_0 lead to more conservative estimates of the effect of Pokémon Go as potential game-related increases in activity would be counted as activity before t_0 (assuming the effect is non-negative).

Measuring Physical Activity

We seek to measure the change in physical activity before and after the time of the first experiential query for Pokémon Go, t_0, when a user presumably started playing the game. We measure physical activity through daily steps as recorded by the 3 axis accelerometer and gyrometer of a wrist-worn consumer activity tracking ftdevice (Microsoft Band). Accelerometer-defined activity measures are preferred over subjective survey-based methods, which have been found to overestimate physical activity by up to 700% [22]. We use steps data from 30 days before the first experiential query (t_0) until 30 days after the first experiential query. We note that, at the time of this study, very few users had been using Pokémon Go for more than 30 days. Furthermore, note that all Pokémon Go users included in our dataset have been using the wearable device for a significant amount of time (median 433 days) such that differences in activity cannot be due to starting to use the wearable device. As not every search engine user whom we identified as a Pokémon Go player is also regularly tracking steps, there are 792 users who tracked steps on at least one day before and after t_0 (see Table 2). We restrict our analysis to these users and compare their activity with the control group described below. Note that the choice of threshold (number of days tracked before and after t_0) does not significantly affect our analysis, as we find very similar results when restricting our analysis to users tracking, for example, 7 days before and after t_0.

Table 2. Number of Pokémon Go users and number of days of steps tracking for these users included in the dataset. We counted days up to 30 days before and after each user’s first experiential query and only considered users with at least one day tracked before and after their first experiential query.

Minimum number of experiential Pokémon Go queries	Number of users	Number of days with steps data
1	792	36,141
2	417	18,804
3	262	11,916
4	199	9132
5	143	6633
6	113	5186
7	85	3819
8	70	3131
9	56	2512
10	50	2218

Control Group

We further compare the differences in activity in the Pokémon Go user population with any changes in the control group, a random sample of US Microsoft wearable users. For example, summertime along with improved weather conditions and potential vacation time might be linked to increases in the steps of the control group as well. As there are no experiential queries for any of the control users, we need to define a suitable substitute for t_0 for the control group in order to compare both groups. We use this reference point t_0 to measure changes in physical activity before and after for both the Pokémon Go user group as well as control users. For the Pokémon Go users, t_0 corresponds to the date of the first experiential query for Pokémon Go (eg, July 6, 2016 or July 7, 2016). One could consider using a single point in time t_0 for all control users, for example the July 6, 2016 release date of Pokémon Go. However, this choice would temporarily align all control users such that weekend, weather, or other effects could lead to confounding. In the Pokémon Go user group, all users have potentially different t_0 based on their first experiential query and therefore such effects are not aligned. In order to match observation periods between both groups, we therefore use the exact same distribution of t_0 for control users; that is, for each control user, we randomly sample a Pokémon Go user and use the same value for t_0 for the control user. This ensures that we will compare physical activity over matching observation periods.

Wear Time

We also measure the wear time of the activity tracking device for each day in the dataset as the number of half hours in a day during which the device was worn on the wrist. Differences in recorded number of steps could potentially stem from simply an increase in wear time rather than an actual increase of physical activity. However, we find that during the study duration the wear time for both Pokémon Go and control users was effectively constant with the ratio between the groups changing by <1% over time. Therefore, we attribute any differences in recorded number of steps to an actual increase in physical activity due to the engagement with Pokémon Go.

Measuring the Impact of Pokémon Go

We measured the impact of Pokémon Go on physical activity by comparing activity levels before and after each user’s first experiential query and relate the level of engagement measured through search queries to the size of the effect on physical activity.

Longitudinal Analysis

We compared the physical activity levels of Pokémon Go users with those of the control group population over time in relation to every user’s first experiential query (t_0). As described above, we used randomly sampled t_0 for users in the control group. We measured the average number of steps over a period of 30 days before the first experiential query until 30 days after the first experiential query. Note that on some days a user might not have recorded any steps and we ignore this user on that day. We measured this average activity separately for the Pokémon Go user group and the control group. To improve graph readability, we smoothed the daily average activity through Gaussian-weighted averaging with a window size of 7 days, but we reported statistical tests on the raw data. We estimated 95% CIs through a bootstrap with 500 resamples [23].

Dose-Response Relationship Between Pokémon Go and Physical Activity

Dose-response relationships between the amount of physical activity and various health outcomes have been well established [4,24,25]. We expect that high engagement with Pokémon Go would be reflected in a larger number of experiential queries. Particularly engaged users might also exhibit larger increases in physical activity. We quantified the exact effect sizes for these increases and studied this potential dose-response relationship between the Pokémon Go–related engagement on a search engine and real-world physical activity. We measured the difference in the average number of daily steps across all users and days for the 30 days before versus 30 days after each user’s first experiential query as the effect size.

Does Everyone Benefit?

We measured the effect on individual user’s physical activity after starting to play Pokémon Go and related the magnitude of this effect to demographic attributes of the user including age, gender, weight status (body mass index, BMI), and prior activity level. We investigated whether only certain user groups are benefiting from the game or whether the potential health benefits might apply more widely to the game's user population. We estimated the effect of playing Pokémon Go on each individual user defined as the difference in the average number of daily steps 30 days before and 30 days after the first experiential query. As this analysis was on user-level, we included only Pokémon Go users with at least seven days of steps tracking before and after this event to reduce noise and applied the same requirement to the control group (we used a threshold of 1 day for analyses on day-level). These constraints resulted in 677 Pokémon Go users and 26,334 control users.

Comparison With Existing Health Apps

We compared the effect of Pokémon Go with the effect of other mobile health apps. The Microsoft Band can be connected to other fitness and health apps and we have data on when these connections first happen (ie, explicit knowledge of t_0 for users of these apps). We studied 4 leading mobile health apps with anonymized names for legal reasons. These apps are regularly rated among the top health apps on both iOS and Android platforms and represent the state-of-the-art in consumer health apps. We also measured the number of daily steps 30 days before a user starts using one of these apps until 30 days after. We only included users who started using the health apps after July 1, 2016 to control for seasonal effects and make the data comparable with our Pokémon Go user group. Again, we restricted our analysis to users who were tracking steps on at least seven days before and after the first experiential query (for Pokémon Go group) or first connecting the health app (for the comparison groups). For the 4 apps, 1155 users were included for app A, 313 for app B, 625 for app C, and 296 users for app D. Note that these users had been using the wearable device for a significant amount of time before connecting to the health app (median time in days for the 4 apps are 87, 57, 103, and 76 days, respectively). Therefore, any differences in average activity were likely due to the connected health app rather than cumulative effects of starting to use a wearable activity tracker. We did not have access to data on the level of engagements with the app.

Estimating the Public Health Impact of Pokémon Go

In order to quantify the effect of Pokémon Go on public health, we estimated (1) how many steps were added to US users' physical activity during the first 30 days, (2) how many users met physical activity guidelines before and after Pokémon Go, and (3) the potential impact on life expectancy if Pokémon Go could sustain the engagement of its users.

The official physical activity guidelines [7,8] are equivalent to approximately 8000 daily steps [26,27]. Only 21% of US adults meet these guidelines. We use all users tracking steps at least seven days before and after their first experiential query for Pokémon Go. We then measured the fraction of users with more than 8000 average daily steps both 30 days before and after the first experiential query. This analysis was repeated for Pokémon Go users with at least one and at least ten experiential queries, and the control group.

If there is a substantial impact on physical activity, Pokémon Go could have a measurable impact on US life expectancy due to well-established health benefits of physical activity on heart disease, stroke, type 2 diabetes, depression, some cancers, obesity, and mortality risk [1,2,3,4,5,6]. If we assume that Pokémon Go users would be able to sustain an activity increase of 1000 daily steps, this would be associated with a 6% lower mortality risk. Using life-table analysis similar to [5] based on mortality risk estimates from [28] and the US 2013 Period Life Table [29], we estimated the impact on life expectancy based on this reduction of mortality risk.

Results

Study Population Demographics

Demographic statistics on identified Pokémon Go users and control users are displayed in Table 3. We found that Pokémon Go users are younger than the average user in our wearable dataset, and much less often female. The gender bias is at least in part due to the gender bias in what users agreed to link data from their Microsoft Band wearables and their online activities to understand product usage and improve Microsoft products. However, there is no evidence for a significantly different effect between males and females in our dataset. Furthermore, there is a significant fraction of overweight and obese users, similar to the proportion expected in the US population [30]. This fraction of overweight and obese users is very similar in the Pokémon Go and control user groups indicating lack of a selection effect based on weight status. The average activity level of Pokémon Go users is below that of the control group indicating that that Pokémon Go is attracting users that get less than average activity. Note that this difference is unlikely to stem from other differences between the 2 groups as younger users are typically more active than older users and males typically get more physical activity than females [12] (ie, we would expect a larger number of steps for the Pokémon Go group given the other differences). Therefore, we do not match the 2 groups on demographic attributes. However, we studied how the effects vary across demographics (see the section "Does Everyone Benefit?").

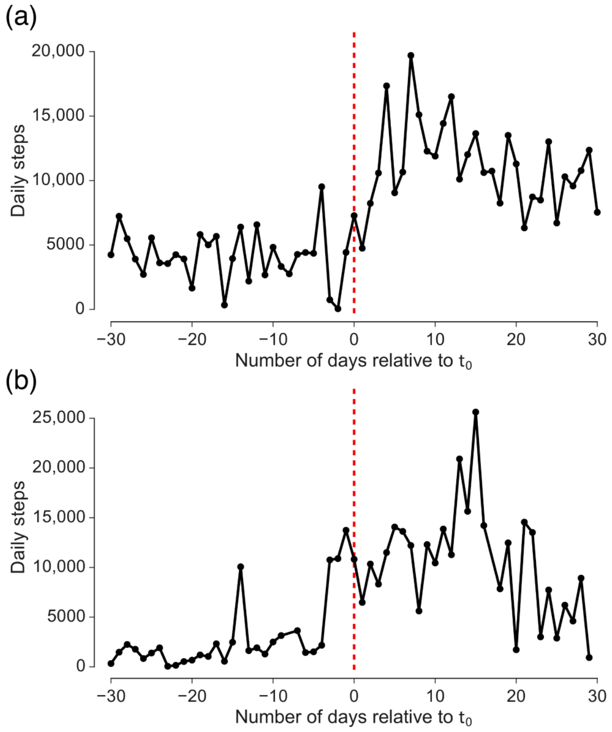
Table 3. Dataset statistics.

Dataset statistic	Pokémon Go users	Wearable users[b]
Number of users	1420	50,000
Number of users with sufficient activity data	792	26,334
Median age	33	42
% female	3.8	25.7
% underweight (BMI[a]<18.5)	1.1	1.2
% normal weight (18.5≤BMI<25)	34.2	31.4
% overweight (25≤BMI < 30)	36.5	38.4
% obese (30≤BMI)	28.2	29.1
Average daily steps overall	6258	6435

[a] BMI: Body Mass Index

[b] Wearable users refers to random sample of US Microsoft Band users. We only consider users with at least one day of steps tracking before and after the user's first experiential query ("sufficient activity data").

Figure 1. Time series of daily steps for 2 sample users (a, b). Both cases show significant increases in daily steps after the first experiential query for Pokémon Go (t_0). Whereas before t_0 both users take less than 5000 steps a day, after t_0 they regularly reach around 15,000 steps a day.



Example Time Series of Physical Activity

Figure 1 displays the daily number of steps before and after the user's first experiential query for 2 example users. Both users significantly increased their activity after their first experiential query for Pokémon Go by several thousand steps each day. We analyzed whether this large increase in physical activity is representative of the study population and how it varied across individuals.

Longitudinal Analysis

Starting to play Pokémon Go is associated with significant increases in physical activity (Figure 2) compared with the control group. The control group slightly decreased their activity by 50 daily steps on average (*P*<.001; we use Mann–Whitney U-Tests for hypothesis tests unless noted otherwise). In contrast, Pokémon Go users increased their activity by 192 daily steps (*P*<.001). The plot shows a steep increase immediately after the first experiential query (t_0), suggesting that the observed increased activity indeed stemmed from engaging with Pokémon Go (*P*<.01, comparing 3 days before and after t_0). We found that Pokémon Go users initially had less activity than the average Microsoft Band user in the United States (dashed blue line; 178 daily steps less; *P*<.001). However, following the start of Pokémon Go play, their activity increased to a level larger than the control group (65 daily steps more; *P*<.001). The activity increase 10 days before t_0 seemed to be statistical noise as we found no evidence for this period to stem from weekend activity or the beginning of summer vacation time due to the variation of t_0 across users.

Activity increases were larger for Pokémon Go users with at least ten experiential queries; that is, users who showed significant interest in Pokémon Go (Figure 2; bottom row). These users were initially significantly less active than the average Microsoft Band user in the United States, getting 5756 daily steps compared with 6435 daily steps in the control group (P<.001). After they started playing Pokémon Go, they exhibited a large increase in activity to an average of 7229 daily steps (1473 daily steps difference; P<.001), which is about 13% larger than the control population (P<.001). This observation suggests that there is a dose-response relationship between interest in Pokémon Go and the effect on physical activity, which we analyze next.

It should be noted that increases in steps before t_0 could stem from starts with the game in advance of queries about Pokémon Go, as we were using the first experiential query as a proxy for the start of play. If users begin to play without ever issuing a search query about Pokémon Go, we could see increases in activity before t_0. However, as we observed steep increases in activity exactly at t_0, this suggests that the proxy for starting is valid for most users.

Note that physical activity for both Pokémon Go user groups (top and bottom rows) decreased again after about 3 to 4 weeks after the first experiential query. However, also note that the activity for the more strongly engaged group (bottom) dropped down to a higher level than they started out with. This suggests that there could be a longer-term behavior change and that future work is needed to study long-term effects of Pokémon Go.

Figure 2. Effect of Pokémon Go on physical activity. Plots show daily steps in absolute numbers for both Pokémon Go users (red) and control users (blue). Plot (a) shows effect for users with at least one experiential query. Plot (b) shows effect for users with at least ten experiential queries. In particular for the users who show significant interest in Pokémon Go (b), we observe large average increases of 1473 steps or 26% over the 30 days following on the first experiential query. Over the same time, the control group (same for both plots) decreased their activity by 50 daily steps on average. Error bars (shaded) in this and all following plots correspond to bootstrapped 95% CIs [23].

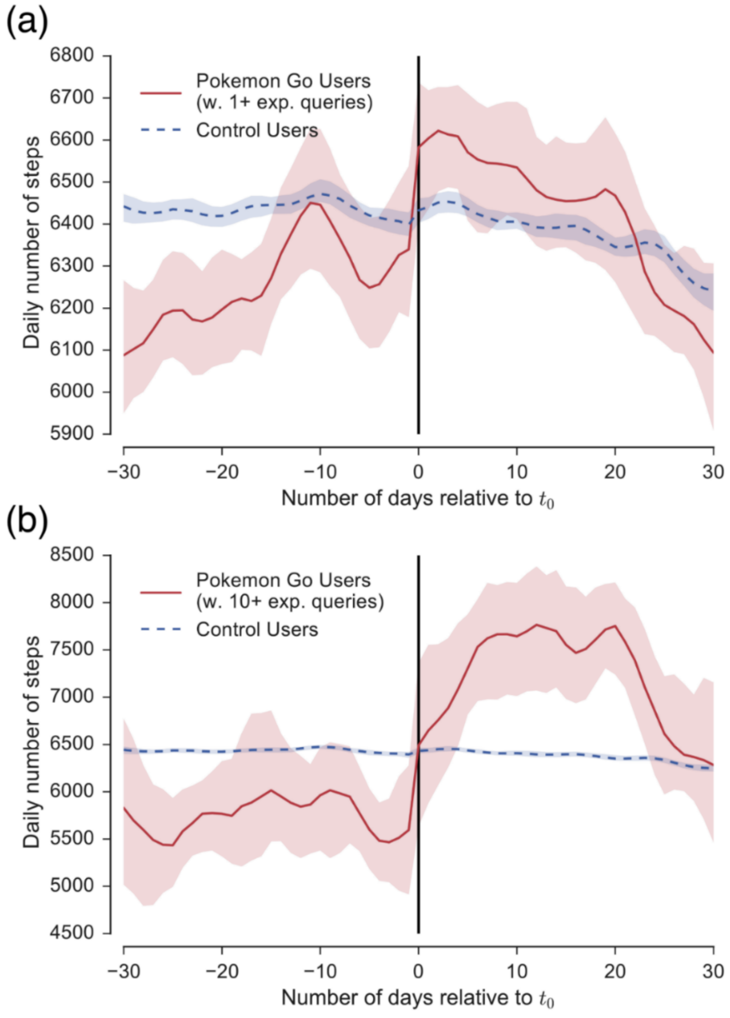
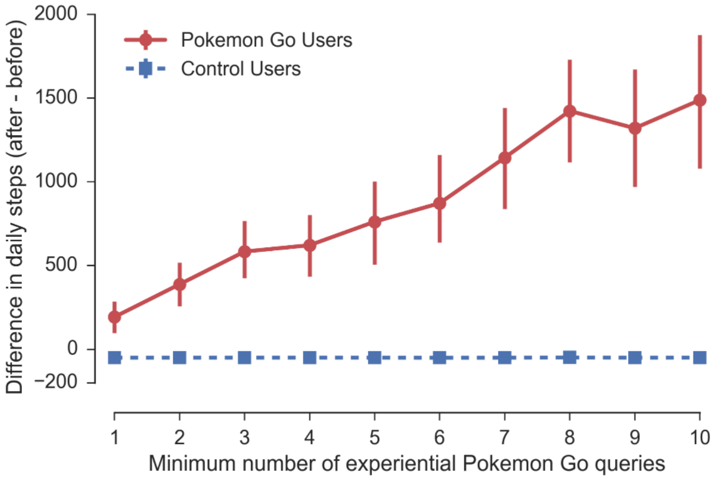


Figure 3. Effect sizes measuring the difference in average number of daily steps between the periods before and after t_0, for different user populations based on the minimum number of experiential queries; that is, toward the right the group of users becomes smaller and increasingly interested in Pokémon Go. At any level, there are significant differences between the effect for Pokémon Go users (red) and the control users (blue). The effect increases linearly with the number of Pokémon Go queries. This dose-response relationship between expressed interest in Pokémon Go and physical activity suggests that these users are in fact playing Pokémon Go and that playing the game makes them more active. CIs for the control group are too small to be visible.



Dose-Response Relationship Between Pokémon Go and Physical Activity

We found a dose-response relationship between the level of engagement with Pokémon Go and the effect on physical activity; that is, users with more experiential queries for Pokémon Go exhibited larger increases in physical activity (see Figure 3). For users who expressed any interest in Pokémon Go, we found significant increases in activity compared with the control group which decreased their activity by 50 steps a day.

Furthermore, we found that these increases in steps scale roughly linearly with the number of experiential queries from 192 daily steps increase (3%) for users with 1 or more experiential queries up to an increase of 1473 daily steps (26%) for users with 10 or more experiential queries. Furthermore, the linear increase in physical activity with the number of experiential Pokémon Go queries suggests that activity increases observed in users querying a search engine for Pokémon Go are causally explained by their engagement with Pokémon Go. If there were other confounding factors that explained the difference in activity between our Pokémon Go group and the control group over time and those changes had nothing to do with Pokémon Go, then one would not expect to find such a clear dose-response relationship as given in Figure 3.

Does Everyone Benefit?

We found that Pokémon Go increased physical activity across men and women of all ages, BMI levels, and prior activity levels (Figure 4). In particular, we found that both Pokémon Go users and control users who were very inactive exhibited large activity increases and users who were relatively active exhibited a decrease in activity on average. However, we found that Pokémon Go users exhibited larger effects than the control across all levels of prior activity (all $P<.025$). We found the largest differences between the 2 groups of users who previously were sedentary, that is, below 5000 daily steps [31]. Furthermore, Pokémon Go users exhibited bigger increases in activity than control users across all age groups (all $P<.040$; except 10-20 year old group which was small), though we found the largest effects for younger users between 10 and 30 years. We also found that the positive effect on physical activity did not vary much across all BMI groups, which is encouraging as obese individuals (30<BMI≤40) were typically less active than healthy subjects [32]. The activity differences in the Pokémon Go groups were always larger than the differences in the control group across all BMI groups (all $P<.021$). Finally, we found that activity differences in the Pokémon Go groups were larger than the differences in the control group for both men and women (all $P<.022$). Increases for women were not significantly different from increases for men ($P=.110$; note small sample size for women).

Figure 4. Effect sizes of physical activity increase or decrease by user demographics, including prior physical activity level (top left), age (top right), body mass index (BMI; bottom left), and gender (bottom right). In all cases, we find that Pokémon Go users (red) exhibit larger changes than their respective control group (blue; see Methods). These results suggest that physical activity increases due to Pokémon Go are not restricted to particular subgroups of users but widely spread across the overall study population.

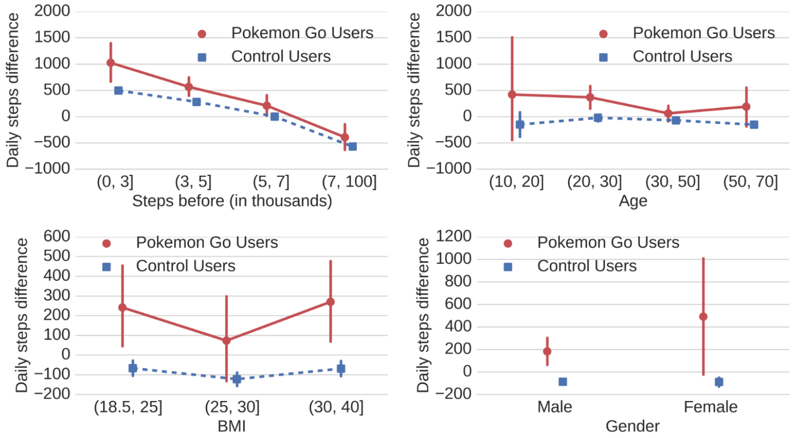
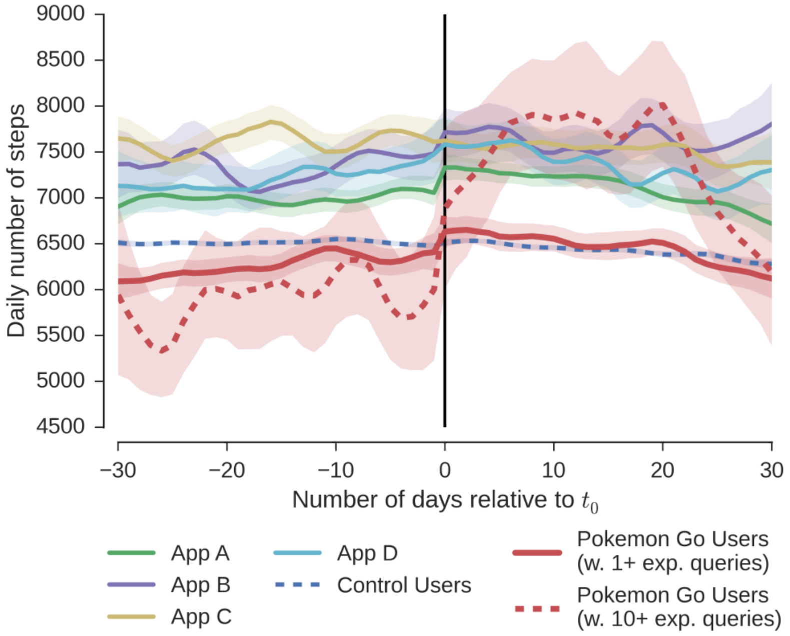


Figure 5. Comparing the effect of the Pokémon Go app with leading consumer health apps (A, B, C, and D). Pokémon Go users are less active than the average wearable user (control) before starting to play, but see larger increases in physical activity compared with the 4 consumer health apps.



Comparison With Existing Health Apps

Pokémon Go led to larger increases in physical activity than other mobile health apps and further attracts more users who are not yet very active. The average daily steps over time is visualized in Figure 5 (using same smoothing method as before). First, we observed that users of all 4 health apps were significantly more active than the average wearable user (6514 daily steps) even before they started to use the health app (6997-7616 daily steps; see activity before t_0 in Figure 5 for apps A, B, C, and D). In contrast, Pokémon Go users were less active than the average user (5901-6265 daily steps). The temporal pattern for the health apps did not contribute strong evidence that these apps were leading to significant behavior change. One exception was app A with its users significantly increasing their activity at day 0. However, this increase in activity was lower compared with the effect of Pokémon Go.

Users of app A increased their activity on average by 111 daily steps or 1.6%. Compare this with 194 (1502) daily steps or 3.1 (25.5%) for Pokémon Go users with at least 1 (10) experiential queries. In particular, users demonstrating large engagement with Pokémon Go exhibited much larger increases in activity than users of any other app in our comparison.

Estimating the Public Health Impact of Pokémon Go

Effect on US Physical Activity

On average, users with one or more experiential query for Pokémon Go increased their physical activity by 192 steps a day for the next 30 days (see results on dose-response relationship). Extrapolating this average effect size to 25 million Pokémon Go users in the United States [16], we found that Pokémon Go added 144 billion steps within the first 30 days to

US physical activity. This is equivalent to walking around the equator 2724 times or 143 round trips to the moon.

Effect on Meeting Activity Guidelines

Using all users tracking steps at least seven days before and after their first experiential query for Pokémon Go, we found that that the fraction of users meeting physical activity guidelines, that is, getting 8000 average daily steps [26,27] stays approximately constant for users with one or more experiential queries (22.0% before vs 21.9% after t_0) and control users (24.1% before vs 23.5%). Note that these estimates are consistent with published estimates in the US adult population (21%) [7,8]. However, for highly engaged Pokémon Go users with at least ten experiential queries, we found that during the 30 days after they started playing, 160% more users achieved 8000 average daily steps (12.2% before vs 31.7% after; relative increase of 160%). For comparison, 21% of US adults met these guidelines [7,8].

Effect on Life Expectancy

We found that more engaged users exhibited average physical activity increases of up to 1473 daily steps. This substantial impact on exercise across the society could have a measurable impact on US life expectancy due to well-established health benefits of physical activity [1,2,3,4,5,6]. If we assume that Pokémon Go users, between 15 and 49 years old, would be able to sustain an activity increase of 1000 daily steps, this would be associated with 41.4 days of additional life expectancy. Across the 25 million US Pokémon Go users [16], this would translate to 2.825 million years additional life added to US users. However, future studies are needed to investigate the long-term effect of Pokémon Go.

Discussion

Principal Findings

The Pokémon Go phenomenon has reached millions of people overnight and dominated news media for weeks after its release [16,17,18,19]. Health professionals have pointed out potential benefits including increased physical activity, spending more time outside, and exploring the neighborhood and city, social interactions, and mastering game challenges, but have also raised concerns such as injury, abduction, trespassing, violence, and cost [15,33]. In this study, we have precisely quantified the impact of Pokémon Go on physical activity and studied the effect on different groups of individuals.

We found that playing the game significantly increased physical activity on the group-level as well as the individual level over an observation period of approximately 4 weeks. The more interest the users showed in Pokémon Go (measured through intensity of search queries seeking details about game usage), the larger the increase in physical activity. For example, users who issued 10 Pokémon Go queries on details of the game within the 2 months after the release of the game increased their activity by 1479 steps a day or 26%.

We found that Pokémon Go increased the activity all across the studied population, largely independent of prior activity level, age, weight status, or gender. These results are encouraging as they suggest that any positive effects due to Pokémon Go are available even to sedentary, obese, and older users. Effectively reaching these users with physical activity interventions is critical for public health [15].

Comparing Pokémon Go with existing mobile health apps, we found further evidence that Pokémon Go is able to reach low activity populations whereas mobile health and fitness apps largely draw from an already active population. Low activity populations would see the largest benefits from improving their activity [7,8]. This highlights the promise of game-based interventions versus traditional approaches, which have often been ineffective for people with low activity levels [13,14].

Given its great popularity, Pokémon Go has significantly affected US physical activity and added an estimated 144 billion steps to US physical activity, which is about 2724 times around the world or 143 round trips to the moon. Furthermore, highly engaged users were almost 3 times as likely to meet official activity guidelines [7,8,26,27] in the 30 days after starting to play Pokémon Go compared with that before. If this user engagement could be sustained, Pokémon Go would have the potential to measurably affect US life expectancy.

These results emphasize the special contribution that activity-encouraging games could have on physical activity and public health. These games attract a wide range of people including those with low prior physical activity. We have demonstrated that such games can lead to significant activity increases and could have a large impact on the society. However, we have also highlighted challenges in realizing this potential. Most importantly, games would need to be able to sustain long-term engagement and lead to sustained behavior change. Furthermore, these games might not be appealing to everyone (eg, we observed males to be more likely to play the games than females), and clearly these games should not replace but complement existing physical activity programs (eg, [2,9,10,11,12,14]). Understanding how to design games and how to bring together games and health interventions will be important to public health in the future. As a first step, our study helps provide guidance on what could come of continuous engagement and with additional engagement.

Limitations

Our study is not without limitations. First, the study population is not a random sample of the US population. Subjects were able to afford a wearable device for activity tracking and willing to share their data for research purposes. Our sample of Pokémon Go users is predominantly male. Furthermore, we use individuals search queries as a proxy for playing Pokémon Go and consider the number of queries as indicating the degree of engagement. However, we find strong evidence that the proxies for usage and engagement are effective. The method identifies a fraction of users that is very similar to independent estimates of Pokémon Go penetration in the United States (see Methods) and we find a strong dose-response relationship between the number of Pokémon Go queries and increased physical activity (see Results). User demographic variables were self-reported. Finally, our follow-up period is currently restricted to 30 days. Future work is needed to study the long-term effectiveness of games such as Pokémon Go to increase physical activity.

Comparison With Prior Work

Few research studies to date have harnessed data obtained from consumer wearable devices to study influences of the devices on physical activity [34]. However, a number of medical studies have examined accelerometer-defined activity [4,31,35], rather than relying on self-report measures. Studies have found that use of pedometers and activity trackers for self-monitoring can help increase activity [36,37], but other studies have reported mixed results [38]. Beyond enabling self-monitoring, encouraging additional activity through reminders lead to increased activity only for the first week after the intervention and did not lead to any significant changes after 6 weeks in a randomized controlled trial [38].

To encourage healthy behavior change, researchers have studied the design of "exergames" [39,40,41], video games combined with exercise activity, and location-based games where game play progresses through a player's location [42]. However, no such game has been nearly as popular and widely used as Pokémon Go. Such games have yet to be integrated into physical activity programs, even though one US college recently announced a physical education class based on Pokémon Go [43]. Recently, a study demonstrated that social influence in online social networks leads to significant increases in average physical activity over a period of several months [44].

There is a growing body of work on using large-scale search query logs to identify subjects with particular conditions for research studies, including such efforts as detecting adverse reactions to medications and identifying signals that could help with screening for cancer [20,45]. Other work has studied activity-related posts on social media to better understand the sharing of health behaviors [46,47,48], but has not yet connected such data to ground-truth health behaviors or focused on interventions on a large scale.

To the best of our knowledge, this is first study of the link between the usage of Pokémon Go or similar games on physical activity and health. Also, this is the first effort to combine data from wearable devices with information drawn from search engine queries.

Conclusions

Novel mobile games that require players to physically move in the real world appear to be an effective complement to traditional physical activity interventions, and they are able to reach millions of engaged users. We studied the effect of Pokémon Go on physical activity through a combination of large-scale wearable sensor data with search engine logs and showed that the game leads to significant increases in physical activity over a period of 30 days, particularly with the engaged users increasing their average activity by 1473 steps a day or 26%. Based on our findings, we estimate that the game has already added an estimated 144 billion steps to US physical activity. If engagement with Pokémon Go could be sustained over the lifetime of its many users, we estimate that the game would add an estimated 2.825 million years of additional lifetime to its US users. However, further studies are needed to investigate potential long-term effects. We see great promise for public health in designing geocentric games like Pokémon Go and in working to sustain users' engagement with them.

Acknowledgments

This research was conducted during an internship at Microsoft Research. All 3 authors conceived the idea for the study and developed the methodology for its execution and were involved in final analysis and writing of the manuscript. The authors thank Jen Hicks for her feedback on the manuscript.

Conflicts of Interest

This work was done while all authors were employees of Microsoft. All work was done as part of the respective authors' research, with no additional or external funding.

References

1. Miles L. Physical activity and health. Nutrition Bulletin 2007;32(4):314-363.
2. Sparling PB, Owen N, Lambert EV, Haskell WL. Promoting physical activity: the new imperative for public health. Health Educ Res 2000 Jun;15(3):367-376 [FREE Full text] [Medline: 10977383]
3. World Health Organization. Global recommendations on physical activity for health. Geneva: World Health Organization; 2010. URL: http://apps.who.int/iris/bitstream/10665/44399/1/9789241599979_eng.pdf [accessed 2016-11-25] [WebCite Cache ID 6mI4Q26mk]
4. Althoff T, Sosic R, Hicks JL, King AC, Delp SL, Leskovec J. Quantifying Dose Response Relationships Between Physical Activity and Health Using Propensity Scores. 2016 Presented at: Neural Information Processing Systems Workshop on Machine Learning for Health (NIPS ML4H); 2016; Barcelona, Spain.
5. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]

6. Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van MW, Lancet Physical Activity Series 2 Executive Committee. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. Lancet 2016 Sep 24;388(10051):1311-1324. [doi: 10.1016/S0140-6736(16)30383-X] [Medline: 27475266]
7. Centers for Disease Control and Prevention. Facts about physical activity. 2014. URL: http://www.cdc.gov/physicalactivity/data/facts.htm [accessed 2016-10-05] [WebCite Cache ID 6l2Glf461]
8. 2008 Physical activity guidelines for Americans. U.S. Department of Health and Human Services (HHS). 2008 Physical activity guidelines for Americans. 2008. URL: https://health.gov/paguidelines/pdf/paguide.pdf [accessed 2016-11-25] [WebCite Cache ID 6mI4okxdN]
9. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC, Lancet Physical Activity Series 2 Executive Committee. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. Lancet 2016 Sep 24;388(10051):1337-1348. [doi: 10.1016/S0140-6736(16)30728-0] [Medline: 27475273]
10. Sallis JF, Bauman A, Pratt M. Environmental and policy interventions to promote physical activity. Am J Prev Med 1998 Nov;15(4):379-397. [Medline: 9838979]
11. Dobbins M, DeCorby K, Robeson P, Husson H, Tirilis D. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6-18. Cochrane Database Syst Rev 2009(1):CD007651. [doi: 10.1002/14651858.CD007651] [Medline: 19160341]
12. Salmon J, Booth ML, Phongsavan P, Murphy N, Timperio A. Promoting physical activity participation among children and adolescents. Epidemiol Rev 2007;29:144-159. [doi: 10.1093/epirev/mxm010] [Medline: 17556765]
13. Dishman RK, Sallis JF, Orenstein DR. The determinants of physical activity and exercise. Public Health Rep 1985;100(2):158-171 [FREE Full text] [Medline: 3920714]
14. Marshall AL. Challenges and opportunities for promoting physical activity in the workplace. J Sci Med Sport 2004 Apr;7(1 Suppl):60-66. [Medline: 15214603]
15. American Heart Assocation. Pokémon Go brings video games outside. 2016. URL: http://news.heart.org/pokemon-go-brings-video-games-outside/ [accessed 2016-10-05] [WebCite Cache ID 6l2HPTpYZ]
16. SurveyMonkey Intelligence Blog. Sonders M. Pokémon GO retention: No, it’s not facing a player loyalty crisis. SurveyMonkey Intelligence Blog. 2016 Oct 05. URL: https://www.surveymonkey.com/business/intelligence/pokemon-go-retention-revenue/ [accessed 2016-10-05] [WebCite Cache ID 6l2HmySIP]
17. Wong J. Pokémon No: game's daily active users, downloads and engagement are down. The Guardian. 2016 Aug 23. URL: https://www.theguardian.com/technology/2016/aug/23/pokemon-go-active-users-down-augmented-reality-games [accessed 2016-10-05] [WebCite Cache ID 6l2FTvvSC]
18. Lynley M. TechCrunch. With 500M downloads, Pokémon Go is coming to the Apple Watch URL: https://techcrunch.com/2016/09/07/pokemon-go-the-hottest-game-on-the-planet-is-coming-to-the-apple-watch/ [WebCite Cache ID 6l2FzwyAD]
19. Perez S. Pokémon Go tops Twitter's daily users, sees more engagement than Facebook. TechCrunch. 2016 Jul 13 URL: https://techcrunch.com/2016/07/13/pokemon-go-tops-twitters-daily-users-sees-more-engagement-than-facebook/ [WebCite Cache ID 6l2Fjj7PB]
20. Paparrizos J, White RW, Horvitz E. Screening for Pancreatic Adenocarcinoma Using Signals From Web Search Logs: Feasibility Study and Results. J Oncol Pract 2016 Jun 7;12(8):737-744. [doi: 10.1200/JOP.2015.010504] [Medline: 27271506]
21. White RW, Wang S, Pant A, Harpaz R, Shukla P, Sun W, et al. Early identification of adverse drug reactions from search log data. J Biomed Inform 2016 Feb;59:42-48. [doi: 10.1016/j.jbi.2015.11.005] [Medline: 26610385]
22. Tucker JM, Welk GJ, Beyler NK. Physical activity in U.S. Adults: Compliance with the Physical Activity Guidelines for Americans. Am J Prev Med 2011 Apr;40(4):454-461. [doi: 10.1016/j.amepre.2010.12.016] [Medline: 21406280]
23. Efron B, Tibshirani RJ. An introduction to the bootstrap. Boca Raton, FL: CRC press; 1994.
24. Dunn AL, Trivedi MH, O'Neal HA. Physical activity dose-response effects on outcomes of depression and anxiety. Med Sci Sports Exerc 2001 Jun;33(6 Suppl):S587-S597. [Medline: 11427783]
25. Lee IM, Skerrett PJ. Physical activity and all-cause mortality: what is the dose-response relation? Med Sci Sports Exerc 2001 Jun;33(6 Suppl):S459-S471. [Medline: 11427772]
26. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De CK, Giles-Corti B, et al. How many steps/day are enough? For adults. Int J Behav Nutr Phys Act 2011;8:79 [FREE Full text] [doi: 10.1186/1479-5868-8-79] [Medline: 21798015]
27. Tudor-Locke C, Leonardi C, Johnson WD, Katzmarzyk PT, Church TS. Accelerometer steps/day translation of moderate-to-vigorous activity. Prev Med 2011;53(1-2):31-33. [doi: 10.1016/j.ypmed.2011.01.014] [Medline: 21295063]
28. Dwyer T, Pezic A, Sun C, Cochrane J, Venn A, Srikanth V, et al. Objectively Measured Daily Steps and Subsequent Long Term All-Cause Mortality: The Tasped Prospective Cohort Study. PLoS One 2015;10(11):e0141274 [FREE Full text] [doi: 10.1371/journal.pone.0141274] [Medline: 26536618]
29. United States Social Security Administration. 2013 Period Life Table. 2013. URL: https://www.ssa.gov/oact/STATS/table4c6.html [accessed 2016-10-05] [WebCite Cache ID 6l2FdRvUU]
30. National Institute of Diabetes and Digestive and Kidney Diseases. Overweight and Obesity Statistics. 2012. URL: https://www.niddk.nih.gov/health-information/health-statistics/Pages/overweight-obesity-statistics.aspx [accessed 2016-10-05] [WebCite Cache ID 6l2FqcgwS]

31. Tudor-Locke C, Bassett DRJ. How many steps/day are enough? Preliminary pedometer indices for public health. Sports Med 2004;34(1):1-8. [Medline: 14715035]
32. Dishman RK, Buckworth J. Increasing physical activity: a quantitative synthesis. Med Sci Sports Exerc 1996 Jun;28(6):706-719. [Medline: 8784759]
33. Serino M, Cordrey K, McLaughlin L, Milanaik RL. Pokémon Go and augmented virtual reality games: a cautionary commentary for parents and pediatricians. Curr Opin Pediatr 2016 Oct;28(5):673-677. [doi: 10.1097/MOP.0000000000000409] [Medline: 27479151]
34. Hayden EC. Mobile-phone health apps deliver data bounty. Nature 2016 Mar 24;531(7595):422-423. [doi: 10.1038/531422a] [Medline: 27008946]
35. 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]
36. Thorup C, Hansen J, Grønkjær M, Andreasen JJ, Nielsen G, Sørensen EE, et al. Cardiac Patients' Walking Activity Determined by a Step Counter in Cardiac Telerehabilitation: Data From the Intervention Arm of a Randomized Controlled Trial. J Med Internet Res 2016;18(4):e69 [FREE Full text] [doi: 10.2196/jmir.5191] [Medline: 27044310]
37. Wang JB, Cataldo JK, Ayala GX, Natarajan L, Cadmus-Bertram LA, White MM, et al. Mobile and Wearable Device Features that Matter in Promoting Physical Activity. J Mob Technol Med 2016 Jul;5(2):2-11. [doi: 10.7309/jmtm.5.2.2] [Medline: 27493694]
38. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable Sensor/Device (Fitbit One) and SMS Text-Messaging Prompts to Increase Physical Activity in Overweight and Obese Adults: A Randomized Controlled Trial. Telemedicine and e-Health 2015 Oct;21(10):782-792. [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
39. Göbel S, Hardy S, Wendel V, Mehm F, Steinmetz R. Serious games for health: Personalized exergames. 2010 Presented at: The 18th ACM international conference on Multimedia; 2010; Firenze, Italy p. 1663-1666.
40. Sinclair J, Hingston P, Masek M. Considerations for the design of exergames. 2007 Presented at: The 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia; 2007; Perth, Australia p. 289-295.
41. Staiano AE, Calvert SL. Exergames for Physical Education Courses: Physical, Social, and Cognitive Benefits. Child Dev Perspect 2011 Jun;5(2):93-98 [FREE Full text] [doi: 10.1111/j.1750-8606.2011.00162.x] [Medline: 22563349]
42. Avouris NM, Yiannoutsou N. A review of mobile location-based games for learning across physical and virtual spaces. J Univers Comput Sci 2012;18(15):2120-2142.
43. Ryssdal K. A college is offering a Pokémon Go class. Marketplace.. 2016 Aug 09. URL: http://www.marketplace.org/2016/08/09/tech/final-note/college-offering-pokemon-go-class [accessed 2016-10-05] [WebCite Cache ID 6l2GdJrU4]
44. Althoff T, Jindal P, Leskovec J. Online Actions with Offline Impact: How Online Social Networks Influence Online and Offline User Behavior. 2017 Presented at: The ACM International Conference on Web Search and Data Mining (WSDM); 2017; Cambridge, United Kingdom.
45. White RW, Horvitz E. From health search to healthcare: explorations of intention and utilization via query logs and user surveys. J Am Med Inform Assoc 2014;21(1):49-55 [FREE Full text] [doi: 10.1136/amiajnl-2012-001473] [Medline: 23666794]
46. Kendall L, Hartzler A, Klasnja P, Pratt W. Descriptive analysis of physical activity conversations on Twitter. 2011 Presented at: In CHI'11 Extended Abstracts on Human Factors in Computing Systems; 2011; Vancouver, BC, Canada p. 1555-1560.
47. Park K, Weber I, Cha M, Lee C. Persistent sharing of fitness app status on Twitter. 2016 Presented at: The 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing; 2016; San Francisco, CA p. 184-194.
48. Teodoro R, Naaman M. Fitter with Twitter: Understanding personal health and fitness activity in social media. 2013 Presented at: Seventh International Conference on Weblogs and Social Media; 2013; Boston, MA.

Abbreviations

BMI: body mass index

Edited by G Eysenbach; submitted 07.10.16; peer-reviewed by I Weber, Y Mejova, A Staiano, C Thorup; comments to author 24.10.16; revised version received 09.11.16; accepted 17.11.16; published 06.12.16.

Please cite as:
Althoff T, White RW, Horvitz E
Influence of Pokémon Go on Physical Activity: Study and Implications
J Med Internet Res 2016;18(12):e315
URL: http://www.jmir.org/2016/12/e315/
doi:10.2196/jmir.6759
PMID:27923778

Original Paper

Short-Term Effectiveness of a Mobile Phone App for Increasing Physical Activity and Adherence to the Mediterranean Diet in Primary Care: A Randomized Controlled Trial (EVIDENT II Study)

Jose I Recio-Rodriguez[1,2], PhD; Cristina Agudo-Conde[1]; Carlos Martin-Cantera[3], MD, PhD; Mª Natividad González-Viejo[4], MD, PhD; Mª Del Carmen Fernandez-Alonso[5], MD, PhD; Maria Soledad Arietaleanizbeaskoa[6]; Yolanda Schmolling-Guinovart[7], MD, PhD; Jose-Angel Maderuelo-Fernandez[1], MD, PhD; Emiliano Rodriguez-Sanchez[1,8], MD, PhD; Manuel A Gomez-Marcos[1,8], MD, PhD; Luis Garcia-Ortiz[1,9], MD, PhD; EVIDENT Investigators[10]

[1]Biomedical Research Institute of Salamanca (IBSAL), Primary Health Care Research Unit, La Alamedilla Health Center, Castilla León Health Service, Salamanca, Spain

[2]University of Salamanca, Department of Nursing and Physiotherapy, Salamanca, Spain

[3]Primary Health Care University Research Institute IDIAP-Jordi Gol, Passeig de Sant Joan Health Center, Catalan Health Service, Department of Medicine. University Autonoma of Barcelona, Barcelona, Spain

[4]Torre Ramona Health Center, Aragon Health Service, Zaragoza, Spain

[5]Casa del Barco Health Center, Castilla y León Health Service, Valladolid, Spain

[6]Primary Health Care Research Unit of Bizkaia.Basque Health Service-Osakidetza, Bilbao, Spain

[7]Río Tajo Health Center, Castilla-La Mancha Health Service, University of Castilla-La Mancha, Talavera de la Reina, Spain

[8]University of Salamanca, Department of Medicine, Salamanca, Spain

[9]University of Salamanca, Department of Biomedical and Diagnostic Sciences, Salamanca, Spain

[10]EVIDENT Investigators. redIAPP: Spanish Research Network for Preventive Activities and Health Promotion in Primary Care, Salamanca, Spain

Corresponding Author:
Luis Garcia-Ortiz, MD, PhD
Biomedical Research Institute of Salamanca (IBSAL)
Primary Health Care Research Unit, La Alamedilla Health Center
Castilla León Health Service
Av. Comuneros 37-31
Salamanca, 37003
Spain
Phone: 34 635542886
Fax: 34 923123644
Email: lgarciao@usal.es

Abstract

Background: The use of mobile phone apps for improving lifestyles has become generalized in the population, although little is still known about their effectiveness in improving health.

Objective: We evaluate the effect of adding an app to standard counseling on increased physical activity (PA) and adherence to the Mediterranean diet, 3 months after implementation.

Methods: A randomized, multicenter clinical trial was carried out. A total of 833 participants were recruited in six primary care centers in Spain through random sampling: 415 in the app+counseling group and 418 in the counseling only group. Counseling on PA and the Mediterranean diet was given to both groups. The app+counseling participants additionally received training in the use of an app designed to promote PA and the Mediterranean diet over a 3-month period. PA was measured with the 7-day Physical Activity Recall (PAR) questionnaire and an accelerometer; adherence to the Mediterranean diet was assessed using the Mediterranean Diet Adherence Screener questionnaire.

Results: Participants were predominantly female in both the app+counseling (249/415, 60.0%) and counseling only (268/418, 64.1%) groups, with a mean age of 51.4 (SD 12.1) and 52.3 (SD 12.0) years, respectively. Leisure-time moderate-to-vigorous physical activity (MVPA) by 7-day PAR increased in the app+counseling (mean 29, 95% CI 5-53 min/week; *P*=.02) but not in the counseling only group (mean 17.4, 95% CI –18 to 53 min/week; *P*=.38). No differences in increase of activity were found between the two groups. The accelerometer recorded a decrease in PA after 3 months in both groups: MVPA mean –55.3 (95% CI –75.8 to –34.9) min/week in app+counseling group and mean –30.1 (95% CI –51.8 to –8.4) min/week in counseling only group. Adherence to the Mediterranean diet increased in both groups (8.4% in app+counseling and 10.4% in counseling only group), with an increase in score of 0.42 and 0.53 points, respectively (*P*<.001), but no difference between groups (*P*=.86).

Conclusions: Leisure-time MVPA increased more in the app+counseling than counseling only group, although no difference was found when comparing the increase between the two groups. Counseling accompanied by printed materials appears to be effective in improving adherence to the Mediterranean diet, although the app does not increase adherence.

ClinicalTrial: Clinicaltrials.gov NCT02016014; https://clinicaltrials.gov/ct2/show/NCT02016014 (Archived by WebCite at http://www.webcitation.org/6mnopADbf)

(J Med Internet Res 2016;18(12):e331) doi:10.2196/jmir.6814

KEYWORDS

physical activity; food; information and communication technologies; arterial aging

Introduction

Regular physical activity (PA) offers considerable physical and psychological health benefits [1], and reduces overall and cardiovascular mortality [2,3] in the general population [4,5]. Despite this fact, most of the population in developed countries do not follow the international recommendations on PA [6,7]. In phase 1 of the “Lifestyles and Arterial Aging” (EVIDENT) trial [8], the proportion of active individuals was very low (31%). Interventions designed to promote PA have revealed a small to moderate effect [9], and could be useful for increasing the number of people adhering to recommended levels of PA. The Mediterranean diet has been shown to be effective in preventing cardiovascular disease [10], and other disorders such as type 2 diabetes [11]. Adherence to the Mediterranean diet is low at only 33%, as documented by the EVIDENT trial [12]. Interventions designed to improve adherence to the Mediterranean diet showed that more frequent contact or intensive interventions appear to be more effective [13].

Information and communication technologies are currently one of the supporting tools that may facilitate such reinforcement and contribute to improving health and changing lifestyles [14]. Many mobile phone apps have been developed with this aim in mind, although the supporting evidence is generally limited [15]. Furthermore, the results are not always uniform, with positive effects in terms of weight loss [16,17], but few or contradictory effects on PA [18,19]. However, few studies have examined effectiveness in large population samples using an app combining PA and food habits.

This study evaluates the short-term (3 months) effects of adding an app in support of standardized counseling to increase PA and adherence to the Mediterranean diet.

Methods

Design

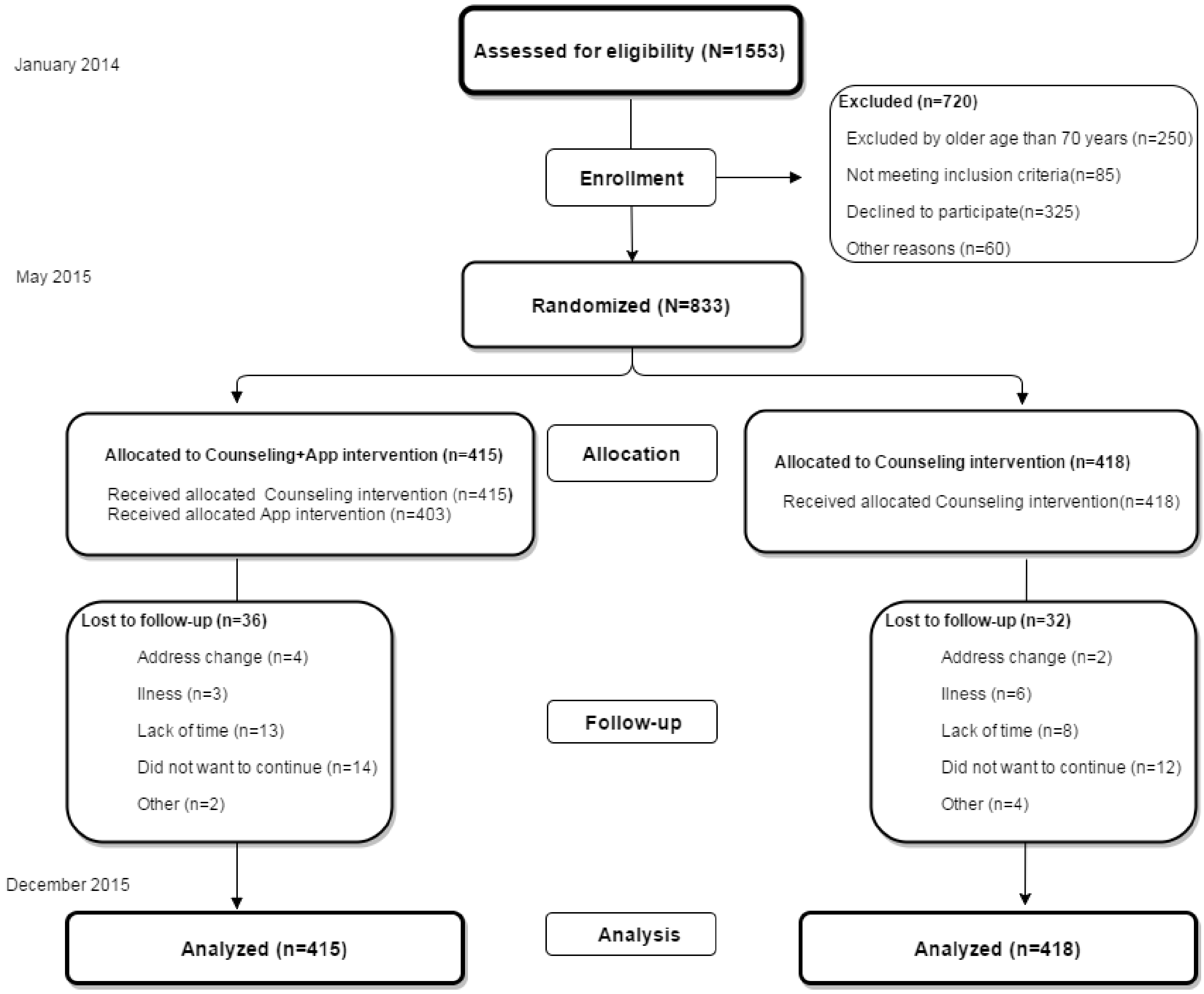
A multicenter randomized controlled trial with two parallel groups was carried out with a follow-up period of 12 months (the EVIDENT II study) [20]. Assessments were made at baseline and after 3 months between January 2014 and December 2015, with evaluation at 12 months in 2016.

Setting and Participants

The study population was selected from the EVIDENT I study [21], including 1553 participants randomly selected from six primary care centers in family practice offices. Participants older than 70 years of age were excluded, as were those unable to do exercise or follow the Mediterranean diet, as well as those individuals meeting any of the exclusion criteria of the EVIDENT I study [20]. Of the participants recruited in the EVIDENT I study, 833 were included in this study (Figure 1) and these participants were randomized in a 1:1 proportion on a centralized basis from Salamanca, using the Epidat 4.0 software package to the counseling+app group (n=415) or the counseling only group (n=418). The investigator who performed the data analysis was blinded. Due to the nature of the study, the participants could not be blinded to the intervention.

The study was approved by the Clinical Research Ethics Committee of the health care area of Salamanca (CEIC de Area de Salud de Salamanca, June 21, 2013) as the reference Committee. In addition, the study was approved by the Ethics Committees of the five collaborating centers (CEIC de Aragón [CEICA], CEIC de IDIAP Jordi Gol, CEIC de Euskadi [CEIC-E], CEIC de Castilla la Mancha, and CEIC de Area de Salud de Valladolid Oeste). All participants signed the informed consent form prior to inclusion in the study, in accordance with the Declaration of Helsinki [22].

Figure 1. Study flowchart: enrollment of participants and completion of study.



Intervention

A research nurse performed a common intervention lasting 30 minutes in both groups. The intervention consisted of standardized counseling in PA and the Mediterranean diet, with the delivery of printed supporting material (leaflet) on the session. The effectiveness of these interventions has been demonstrated in the "Experimental Program for Physical Activity Promotion" (PEPAF) study [23] and the "Prevención con Dieta Mediterránea" (PREDIMED) study [10].

For the counseling+app group , the participants also received training in the use of a mobile phone app designed to promote the Mediterranean diet and increase PA over a 3-month period. The app was designed by software engineers in collaboration with dietitians and PA experts, with an easy-to-use interface for logging food and exercise. A 15-minute initial visit was used to provide training in the use of the device, which was employed daily for the full 3-month period of the intervention. The participants were instructed to enter their food intake (breakfast, lunch, afternoon snack, and dinner). Based on adequate proportions of macronutrients (carbohydrates: 50%-60%; protein: 10%-15%; fats: 30%-35%; fiber: >22 g/day), following recommendations of the of the , National Academy of Sciences (US Department of Health and Human Services), a personalized recommendation was provided. Regular PA was recorded with the accelerometer of the device, together with user input of activities performed without the mobile phone (eg, swimming, football). Lastly, at the end of the day, the app reported a summary of food intake and PA performed and a balance of ingested and spent calories. The app, in turn, generated a recommended plan for the following days, with a view to improve eating habits and increase PA. A new visit took place 1 week after supplying the device to confirm that it was being used correctly. The mobile phone was returned after 3 months, coinciding with the common follow-up visit. Adherence to the mobile phone app was assessed by the number of days of recordings in the device.

Outcome Measures and Follow-Up

The main outcome measures were change in PA and adherence to the Mediterranean diet at 3 months in the app+counseling group compared to the counseling only group. Other outcome measures were also collected, including blood pressure, waist circumference, body mass index (BMI), and laboratory parameters. A detailed description has been published elsewhere of how clinical data were collected [20].

Physical Activity

Physical activity was measured with an accelerometer and the 7-day Physical Activity Recall (7-day PAR) questionnaire. The ActiGraph GT3X accelerometer (ActiGraph, Shalimar, FL, USA) was used and has been previously validated [24-26]. Activity “counts” were recorded to the internal memory of the accelerometers by converting acceleration units over a given epoch [27]. Participants wore the accelerometer fastened with an elastic strap to the right side of the waist for seven consecutive days during habitual PA, except for bathing and performing activities in the water. The accelerometer was set to record PA data every minute. Inclusion criteria were a minimum of 5 days of recording, including at least one weekend day and at least 600 registered minutes per day. The first and last day’s data were excluded to analyze full days only, and the uptime was adjusted to 7 days. The main outcome variable from the activity monitor was the mean intensity of PA (counts/minute). The intensity of PA was rated according to the cut-off points proposed by Freedson [28].

The 7-day PAR is a semistructured interview (10-15 minutes) in which participants provide an estimate of the number of hours dedicated to physical or occupational activities requiring at least a moderate effort in the past 7 days. The categories collected are moderate, vigorous, and very vigorous PA. The amount of time dedicated to each activity was multiplied by the mean metabolic equivalent (MET) of each category: light activity=1.5, moderate activity=4, vigorous activity=6, and very vigorous activity=10. The dose of physical exercise was estimated in MET-minutes/week. Active individuals were considered as those doing at least 30 minutes of moderate activity 5 days per week, or at least 20 minutes of vigorous activity 3 days per week [29].

Nutrition

Adherence to the Mediterranean diet, as a nutrition primary endpoint, was measured using the validated 14-point Mediterranean Diet Adherence Screener (MEDAS) [30], developed by the PREDIMED study group. The 14-item screener includes 12 questions on food consumption frequency and two questions on food intake habits considered characteristic of the Spanish Mediterranean diet. Each question was scored as 0 or 1, and the total score ranged from 0 to 14. Adequate adherence to the Mediterranean diet was assumed when the total score was ≥9 points.

Statistical Analysis

Estimation of sample size was made for the main study endpoints. Assuming alpha=.05 and beta=.20, with a SD 154 counts/minute, we would need 828 participants (414 per group) to detect an increase of 30 counts/minute in the app+counseling versus counseling only group; for a SD of 2 points in Mediterranean diet, we would need 504 participants (252 per group) to detect an increase of 0.5 points. We considered it sufficient to include 833 participants in order to detect clinically relevant differences in the main study endpoints.

The results were expressed as mean and standard deviation for quantitative variables and as the frequency distribution for qualitative variables. Analysis of the results was made on an intention-to-treat basis. The chi-square test and Fisher test were used to analyze the association between independent qualitative variables. The Student *t* test was used for the comparison of means between two groups and the paired *t* test was applied to assess changes within the same group. Analysis of variance (ANOVA) was used for comparison of means between more than two groups. In order to analyze the effect of the intervention, we compared the changes observed between the counseling only and app+counseling groups by ANCOVA test, adjusting for baseline measures of each variable. To evaluate the effect of adherence to the app in the increase of PA measured by the accelerometer, we performed a multivariate analysis based on the general lineal model, adjusting the results for age and sex. The intraclass correlation coefficients of outcome data for centers were for MET-minutes/week in leisure time (7 day PAR; ρ=.009), for total moderate-to-vigorous PA (MVPA) time/week (accelerometer; ρ=.002), and for the score of adherence to Mediterranean diet (MEDAS; ρ=.009). The contrasting of hypotheses established alpha=.05. The data were analyzed using SPSS version 23.0 (IBM Corp, Armonk, NY, USA). A value of *P*<.05 was considered statistically significant.

Results

Baseline Characteristics of the Participants and Follow-Up

The 833 participants were predominantly females in both the app+counseling (249/415, 60.0%) and counseling only (268/418, 64.1%) groups, with a mean age of 51.4 (SD 12.1) and 52.3 (SD 12.0) years, respectively. Likewise, no differences were observed between the two groups in terms of the other demographic and clinical characteristics (Table 1). Regarding PA evaluated with the 7-day PAR, we found no significant difference between the groups; app+counseling reached a mean 864.6 (SD 1407.7) MET-minutes/week, whereas the counseling only group reached a mean 865.8 (SD 1330.6) MET-minutes/week (*P*=.34). For the parameters analyzed with the accelerometer, the results were also similar in both groups (Table 2). For adherence to the Mediterranean diet, the mean score was 7.6 (SD 2.1) in the app+counseling group and mean 7.4 (SD 1.9) in the counseling only group, with an adequate adherence rate of 34.2% (142/315) in the app+counseling group versus 28.5% (119/318) in the counseling only group (*P*=.09) (Table 2 and Multimedia Appendix 1).

Table 1. Baseline characteristics of the study population (N=833).

Variable	App+counseling (n=415)	Counseling only (n=418)	*P* value
Age (years), mean (SD)	51.4 (12.1)	52.3 (12.0)	.28
Gender (female), n (%)	249 (60.0)	268 (64.1)	.23
Work situation, n (%)			.25
Works outside home	228 (54.9)	203 (48.6)	
Homemaker	53 (12.8)	72 (17.2)	
Retired	77 (18.6)	89 (21.3)	
Student	10 (2.4)	8 (1.9)	
Unemployed	47 (11.3)	46 (11.0)	
Educational level, n (%)			
University studies	117 (28.2)	132 (31.6)	.42
Middle or high school	208 (50.1)	208 (49.8)	
Elementary school	90 (21.7)	78 (18.7)	
Smoking, n (%)			.20
Nonsmoker	190 (45.8)	166 (39.7)	
Smoker	94 (22.7)	108 (25.8)	
Former smoker	131 (31.6)	144 (34.4)	
Waist circumference (cm), mean (SD)	95.2 (13.2)	94.8 (13.1)	.71
BMI (kg/m^2), mean (SD)	28.1 (5.1)	27.6 (4.59)	.14
BMI category, n (%)			.50
<25	117 (28.2)	131 (31.3)	
25-30	172 (41.4)	173 (41.4)	
>30	126 (30.4)	114 (27.3)	
Systolic blood pressure (mmHg), mean (SD)	124 (16)	124 (17)	.75
Diastolic blood pressure (mmHg), mean (SD)	76 (10)	76 (10)	.41
Total cholesterol (mg/dL), mean (SD)	202 (35)	206 (37)	.08
Triglycerides (mg/dL), mean (SD)	112 (63)	107 (67)	.29
Glycated hemoglobin (%), mean (SD)	5.5 (0.5)	5.5 (0.6)	.87
Hypertension, n (%)	145 (34.9)	133 (31.8)	.34
Dyslipidemia, n (%)	116 (28.2)	113 (27.3)	.77
Diabetes, n (%)	32 (7.7)	30 (7.2)	.77
Medication use, n (%)			
Antihypertensive drugs	108 (26.0)	95 (22.7)	.29
Lipid-lowering drugs	90 (21.7)	74 (17.7)	.16
Antidiabetic drugs	24 (5.8)	28 (6.7)	.67
Physical activity stage of change, n (%)[a]			.35
Precontemplation	57 (14.0)	73 (17.7)	
Contemplation	28 (6.9)	38 (9.2)	
Preparation	58 (14.2)	51 (12.4)	
Action	30 (7.4)	26 (6.3)	
Maintenance	235 (57.6)	224 (54.4)	
Dietary habits stage of change, n (%)			.91

Variable	App+counseling (n=415)	Counseling only (n=418)	*P* value
Precontemplation	34 (8.3)	34 (8.2)	
Contemplation	26 (6.3)	20 (4.8)	
Preparation	59 (14.3)	57 (13.8)	
Action	18 (4.4)	18 (4.3)	
Maintenance	275 (66.7)	285 (68.8)	
Participants with total Mediterranean diet score ≥9 points, n (%)	142 (34.2)	119 (28.5)	.09
Score for adherence to Mediterranean diet, mean (SD)	7.6 (2.1)	7.4 (2.0)	.09

[a] Stage of change by Prochaska and Diclemente model.

Table 2. Baseline physical activity by 7-day Physical Activity Recall (PAR) questionnaire and accelerometer data (N=833).

Measures of physical activity	App+counseling (n=415)	Counseling only (n=418)	*P* value
7-day PAR[a]			
Total minutes moderate activity, mean (SD)	152.7 (264.8)	154.9 (258.2)	.90
Total minutes moderate activity in leisure time, mean (SD)	131.2 (213.1)	148.0 (249.4)	.29
Total minutes vigorous/very vigorous activity, mean (SD)	29.9 (99.2)	30.0 (106.2)	.98
Total minutes vigorous/very vigorous activity in leisure time, mean (SD)	28.0 (97.1)	27.9 (98.2)	.99
Total minutes moderate vigorous/very vigorous activity, mean (SD)	182.6 (293.0)	184.9 (284.7)	.91
Total minutes moderate vigorous/very vigorous activity in leisure time, mean (SD)	159.1 (228.9)	175.9 (271.6)	.33
MET-minutes/week, mean (SD)	864.6 (1407.8)	865.8 (1330.6)	.99
MET-minutes/week in leisure time, mean (SD)	764.4 (1119.7)	825.6 (1263.3)	.46
Active,[b] n (%)	114 (27.5)	118 (28.2)	.82
Accelerometer			
Steps/day, mean (SD)	9992.3 (3847.3)	9708.3 (3930.9)	.31
Counts minute/week, mean (SD)	69.0 (70.4)	65.9 (69.4)	.54
Sedentary minutes/week, mean (SD)	8327.0 (540.4)	8341.4 (526.0)	.71
Light minutes/week, mean (SD)	1298.3 (436.9)	1307.2 (423.2)	.77
Moderate minutes/week, mean (SD)	438.0 (205.4)	413.3 (212.6)	.10
Vigorous/very vigorous minutes/week, mean (SD)	16.7 (38.9)	18.2 (45.8)	.62
Total MVPA[c] minutes/week, mean (SD)	455.4 (215.9)	432.7 (222.5)	.15
MET-minutes/week, mean (SD)	1850.8 (891.7)	1762.6 (922.0)	.18
>450 MET-minutes/week, n (%)	368 (96.6)	373 (94.4)	.15

[a] MET: metabolic equivalent

[b] Active were considered as those doing at least 30 minutes of moderate activity for 5 days a week, or at least 20 minutes of vigorous activity for 3 days a week.

[c] MVPA: moderate-to-vigorous physical activity.

Of the 833 participants included in the study, 36 of 415 (8.6%) in the app+counseling group were lost at the 3-month time point versus 32 of 418 (7.6%) in the counseling only group (Figure 1).

Changes in Physical Activity and Adherence to the Mediterranean Diet

Based on data from the 7-day PAR, both groups increased their PA after 3 months, although only the app+counseling group reached statistical significance for the criteria leisure-time moderate activity (mean 28, 95% CI 6-50 minutes/week) and leisure-time MVPA (mean 29, 95% CI 5-53 minutes/week). Although the increase in activity in the app+counseling group was greater than in the counseling only group for all analyzed variables, no significant differences were observed between the two groups (Table 3). In relation to PA evaluated with the accelerometer, we recorded a decrease in daily steps, counts/minute, and time spent at the different levels of activity (except for vigorous/very vigorous activity) with an increase in sedentary time in both groups. Furthermore, no differences were observed when comparing the changes between the two groups (Table 2). Both groups increased adherence to the Mediterranean diet to a similar degree after 3 months versus baseline (app+counseling: 8.4%; counseling only: 10.4%), with an increase in overall score of mean 0.42 (95% CI 0.24-0.60) points in the app+counseling group and mean 0.53 (95% CI 0.35-0.71) points in the counseling only group (Table 4).

Table 3. Changes in physical activity and sedentary lifestyle at 3 months compared to baseline.

Measures of physical activity	App+counseling (n=379)		Counseling group (n=386)		Mean difference (app+counseling—counseling group)[d]	
	Mean (95% CI)	*P*	Mean (95% CI)	*P*	Mean (95% CI)	*P*
7-day PAR[a]						
Total minutes moderate activity/week	20.6 (–8.1, 49.2)	.16	8.3 (–18.6, 35.2)	.54	10.7 (–23.8, 45.2)	.54
Minutes moderate activity in leisure minutes/week	28.4 (6.0, 50.8)	.01	12.8 (–13.2, 38.8)	.33	7.8 (–23.5, 39.2)	.62
Total minutes vigorous/very vigorous activity/week	2.8 (–7.6, 13.1)	.60	–0.7 (–10.0, 8.7)	.89	3.5 (–8.3, 15.2)	.56
Minutes vigorous/very vigorous activity in leisure minutes/week	0.7 (–9.5, 10.9)	.89	–1.1 (–10.3, 8.1)	.82	1.9 (–9.1, 13.0)	.73
Total minutes MVPA[c]/week	23.3 (–5.4, 52.1)	.11	7.7 (–19.8, 35.2)	.58	14.2 (–20.1, 48.5)	.42
Minutes MVPA in leisure minutes/week	29.1 (4.9, 53.3)	.02	11.7 (–14.6, 38.1)	.38	9.5 (–22.6, 41.6)	.56
MET[b]-minutes/week	88.8 (–42.8, 220.3)	.18	14.5 (–108.7, 137.8)	.82	72.4 (–76.4, 221.2)	.34
MET-minutes/week in leisure time	110.5 (–5.0, 225.9)	.06	26.9 (–90.3, 144.2)	.65	51.5 (–90.0, 192.9)	.47
Accelerometer[e]						
Steps/day	–1042.1 (–1401.7, –682.6)	<.001	–584.2 (–961.2, –207.1)	.002	–354.6 (–833.7, 124.4)	.15
Counts minutes/week	–12.9 (–18.6, –7.3)	<.001	–6.8 (–13.3, –0.3)	.04	–5.1 (–12.7, 2.5)	.19
Sedentary minutes/week	167.7 (114.9, 220.5)	<.001	125.6 (73.7, 177.6)	<.001	32.9 (–37.6, 103.3)	.36
Light minutes/week	–113.0 –(154.4, –71.6)	<.001	–96.6 (–137.3, –55.8)	<.001	–14.3 (–69.6, 41.1)	.61
Moderate minute /week	–51.3 (–71.3, –31.4)	<.001	–26.3 (–47.0, –5.5)	.01	–14.8 (–40.8, 11.1)	.26
Vigorous very vigorous minutes/week	–3.4 (–6.9, , 0.2)	.06	–2.8 (–6.7, 1.0)	.15	–1.1 (–5.6, 3.4)	.62
Total MVPA minutes/week	–55.3 (–75.8, –34.9)	<.001	–30.1 (–51.8, –8.4)	.01	–15.9 (–42.8, 10.9)	.24
MET-minutes/week	–229.3 (–313.2, –145.4)	<.001	–118.6 (–208.6, –28.7)	.01	74.9 (–186.1, 36.2)	.19

[a] 7-day PAR:7-day Physical Activity Recall;

[b] MET: metabolic equivalent

[c] MVPA: moderate-to-vigorous physical activity.

[d] Changes in app+counseling and counseling only group=data at 3 months–baseline.

[e] In accelerometer measurement, 335 participants in app+counseling group and 344 in counseling only group.

Table 4. Changes in the Mediterranean diet at 3 months compared to baseline.

Criteria Mediterranean diet	App+counseling[a] (n=379)		Counseling group[a] (n=386)		Mean difference (app+counseling–counseling)	
	% Mean (95% CI)	*P*	% Mean (95% CI)	*P*	% Mean (95% CI)	*P*
Using olive oil as the principal source of fat for cooking	3.2 (0.9, 5.5)	.01	2.6 (0.4, 4.8)	.02	0.6 (–2.6, 3.7)	.72
≥4 tbsp (54 g) of olive oil/day (eg, used in frying, salads, meals eaten away from home)	0.5 (–4.4, 5.5)	.83	2.1 (–2.8, 6.9)	.40	–1.6 (–8.5, 5.4)	.66
<.0017.6 (2.9, 12.2).0010.4 (–6.0, 6.8).902 or more servings of vegetables/day	8.2 (3.1, 13.3)	.002	12.3 (7.2, 17.4)	<.001	–4.1 (–11.3, 3.2)	.27
1 serving of red meat or sausage/day8.0 (3.5, 12.4)	1.6 (–2.6, 5.8)	.45	3.6 (0.1, 7.2)	.04	–2.1 (–7.5, 3.4)	.46
3 or more pieces of fruit/day						
1 serving of animal fat/day	0.8 (–2.2, 3.8)	.60	0.8 (–2.1, 3.7)	.59	0.0 (–4.1, 4.1)	.99
1 cup (100 mL) of sugar-sweetened beverages/day	2.1 (–1.2, 5.4)	.21	1.3 (–2.1, 4.8)	.46	0.8 (–4.0, 5.6)	.74
≥7 servings of red wine/week	–0.3 (–3.4, 2.9)	.87	0.3 (–2.1, 2.6)	.83	–0.5 (–4.5, 3.4)	.79
≥3 servings of legumes/week	–3.4 (–8.1, 1.2)	.14	–1.6 (–5.5, 2.3)	.43	–1.9 (–7.9, 4.1)	.54
≥3 servings of fish/week	5.3 (0.5, 10.1)	.03	3.9 (–0.8, 8.6)	.10 1.4 (–5.3, 8.1)	.68	
<2 commercial pastries/week	6.9 (1.5, 12.2)	.01	8.6 (3.6, 13.7)	.001	–1.7 (–9.1, 5.6)	.64
≥3 servings of nuts/week	2.1 (–2.9, 7.1)	.40	4.7 (–0.2, 9.5)	.05	–2.6 (–9.5, 4.4)	.47
Preferring white meat over red meat?	6.4 (1.4, 11.3)	.01	10.4 (6.3, 14.6)	<.001	–4.1 (–10.5, 2.4)	.22
≥2 servings/week of a dish with a traditional sauce of tomatoes, garlic, onion, or leeks sautéed in olive oil	3.4 (–1.8, 8.7)	.20	–0.5 (–5.9, 4.9)	.85	4.0 (–3.6, 11.5)	.30
Study participants with a total score ≥9 points	8.4 (3.4, 13.5)	.001	10.4 (5.0, 15.8)	<.001	–1.9 (–9.3, 5.5)	.61
Difference Mediterranean diet score	0.42 (0.24, 0.60)	<.001	0.53 (0.35, 0.71)	<.001	–0.02 (–0.25, 0.21)	.86

[a] Changes in app+counseling and counseling only group=data at 3 months–baseline.

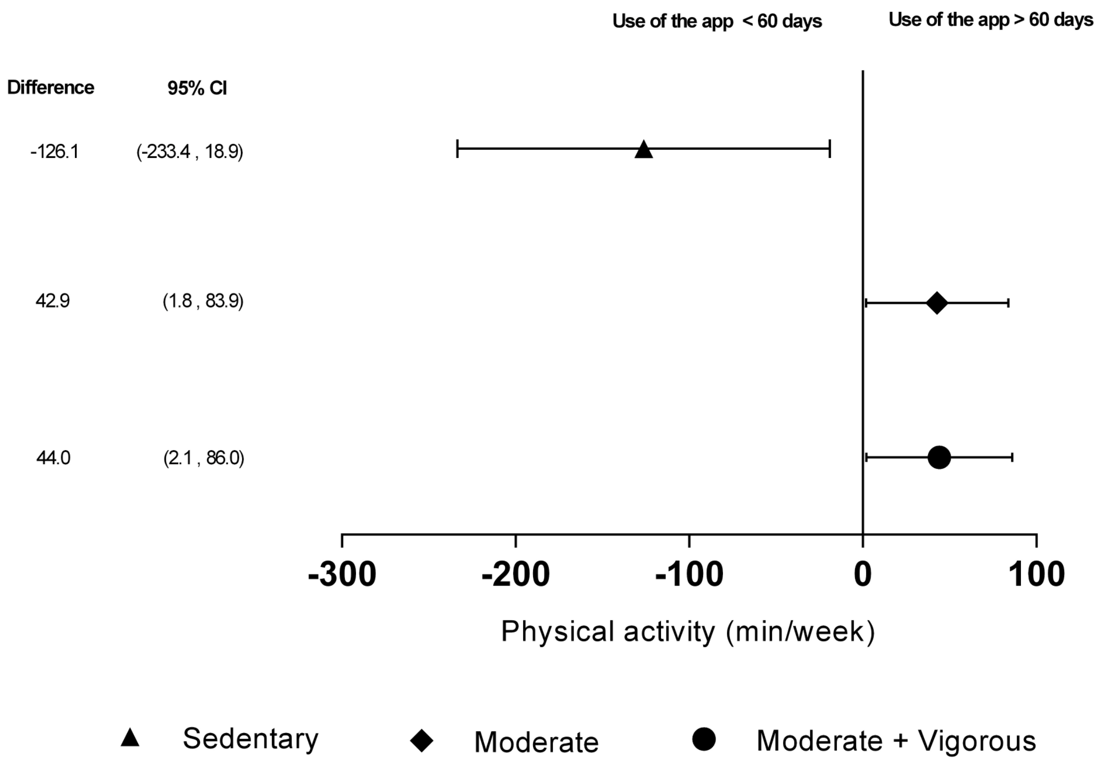
Adherence to the Mobile Phone App

In the app+counseling group, 56.8% (236/315) of participants used the mobile phone for more than 60 days. Although there was a decrease in PA after 3 months, as evaluated by the accelerometer in both groups, the decrease was less pronounced in the group that used the mobile phone most (>60 days). This group showed a net increase in the time of moderate PA of mean 42.9 (95%CI 1.8-83.9) minutes/week and MVPA of mean 44.0 (95% CI 2.1-86.0) minutes/week, and a net decrease in sedentary time of mean 126.1 (95% CI 18.9-233.4) minutes/week. We also recorded a net increase in counts/minute of mean 766.6 (95% CI 26.2-1506.9) (Figure 2).

Analysis by Subgroups

In the male subgroup, we observed a greater decrease in MVPA by accelerometry in the app+counseling than counseling only group (Multimedia Appendixes 2 and 3). In the female subgroup, we observed an increase in leisure-time moderate physical activity as evaluated with the 7-day PAR in the app+counseling versus counseling only group, although the statistical significance was lost after adjusting for baseline data (Multimedia Appendix 4).

Figure 2. Adherence to the mobile phone app. Changes in physical activity evaluated with the accelerometer, according to adherence to the mobile phone app (number of days with a record in the app). Higher adherence (>60 days): 56.8%, 236/415; lower adherence (≤60 days): 43.1%, 179/415. (0-6 days: 17.8%, 74/415; 7-30 days: 10.3%, 43/415; 31-60 days: 14.9%, 62/415).



Discussion

Principal Findings

Although some randomized controlled clinical trials have analyzed the effect of mobile phone apps in promoting healthy lifestyles, the EVIDENT II trial has included the largest number of participants (N=833) and also has the longest follow-up (12 months). The main findings at short-term follow-up (3 months) were an increase in PA as evaluated by the 7-day PAR in both groups (although greater in app+counseling), and an increase in the time dedicated to leisure-time moderate/vigorous activities. However, assessment with the accelerometer revealed a similar decrease in PA in both groups. Adherence to the Mediterranean diet was seen to increase in both groups, as evidenced by the MEDAS score. Lastly, in the accelerometer analysis, the participants in the app+counseling group that most used the app showed a net increase in MVPA time and a net decrease in sedentary time.

Comparison With Prior Work

At this time, there is still no conclusive evidence of the effectiveness of apps for mobile phones in improving lifestyles. In this regard, the meta-analysis published by Flores et al [16] found that interventions with apps had some impact in terms of weight loss (mean BMI 0.43 kg/m^2), although no improvement in terms of increased PA was observed. Partridge et al [31], in a sample of 250 participants between ages 18 and 35 years, evaluated PA using the International Physical Activity Questionnaire (IPAQ), with results similar to those obtained in our study with the 7-day PAR. In both cases, there was a greater increase in PA in the intervention group versus the control group, although the differences between them were not significant. In turn, Laing et al [17], in a randomized controlled study of 212 overweight individuals with a mean age of 43 years, found the use of an app (MyFitnessPal) had no impact on either weight loss or increased PA as assessed by questionnaire.

However, the SMART MOVE study [32], which had 90 participants (45 in each group), PA was assessed from the steps estimated by the mobile phone pedometer and an increase was recorded after 8 weeks in the intervention group (1631 steps), whereas a decrease was observed in the control group (–386 steps). The baseline values in the two groups were 4365 and 5138 steps/day, respectively. In the baseline evaluation of the EVIDENT II study, the mean number of steps/day as determined with the accelerometer was 9992 in the app+counseling group and 9708 in the counseling only group. This was followed by a decrease in both groups after 3 months, possibly because the baseline values were very high. We have found no studies involving accelerometer interventions in adults with the published data being limited to younger participants. Direito et al [18] compared two intervention groups using two different apps with a control group. Physical activity, evaluated with an accelerometer, decreased in both the control group and in one of the intervention groups after 8 weeks coinciding with the findings of the EVIDENT study, which saw practically no changes in the data assessed with the Physical Activity Questionnaire for Adolescents (PAQ-A). The decrease in accelerometer recordings is probably attributable to a Hawthorne effect associated with utilization of the device—the increase in usual activity being more evident at baseline than after 3 months due to a certain loss of effect. This circumstance could limit the usefulness of the accelerometer in evaluating the effect of the interventions, despite the method being objective.

In the EVIDENT trial, nutritional counseling was seen to increase the overall score of adherence to the Mediterranean diet. Counseling was standardized in both groups with all participants receiving an informative leaflet [20]. This type of nutritional counseling has shown improvements in food habits, with a moderate increase in the consumption of fruit, vegetables, and fiber, especially when written materials are supplied in support of counseling [13]. However, the added use of an app did not result in significant differences between the overall groups or subgroups. There is little evidence of the effectiveness of apps in improving food habits and, in general, the results obtained are modest and come from studies with small sample sizes [33].

Nevertheless, the use of new technologies achieved some change in a study of young individuals aged between 18 and 35 years, with a slight increase in vegetable intake and a decrease in the consumption of sugared beverages [31]. Furthermore, a lesser calorie and fat intake was recorded, resulting in increased weight loss [33]. On the other hand, Coughlin et al [33] considered that heterogeneity in the functional characteristics of the different apps makes it more difficult to draw conclusions and to estimate the magnitude of their effect. In turn, Wang et al [34] suggested that effectiveness can be increased by orienting these tools toward personalized needs, such as self-education and the gaining of awareness of personal food intake. In this regard, one of the novelties of the EVIDENT app is the incorporation of weekly notifications on the benefits and characteristics of the consumption of vegetables, fruit, olive oil, fish, and tomato sauce prepared with vegetables and olive oil—all being traditional ingredients of the Mediterranean diet.

Limitations

Our study also has several limitations. The nature of the intervention precludes blinding of the participants and this could influence the results obtained. Also, the main findings of the study are based on self-reported information for adherence to both the Mediterranean diet and to PA. Lastly, the recorded loss rate of close to 10% may have biased the study sample composition to some extent because certain populations may have experienced difficulties using the app and consequently decided to leave the study. During the study period, although we recommended not using other apps or wearables that register nutrition and PA, we have no total guarantee that any participant did not use them.

Conclusions

Physical activity, evaluated with the 7-day PAR, increased more in the app+counseling group than the counseling only group for leisure-time MVPA, although no difference was found when comparing the increase between the two groups. Improved adherence to the app appears to be associated with better results in terms of PA evaluated with the accelerometer. Counseling accompanied by printed materials appears to be effective in improving adherence to the Mediterranean diet, although the app for mobile phones did not increase effectiveness. Further studies are needed to determine which population subgroups may benefit most from interventions based on information and communication technologies.

Acknowledgments

EVIDENT Investigators Group

1. Coordinating center: La Alamedilla Health Center (Health Service of Castilla y León): Luis Garcia-Ortiz, Jose I Recio-Rodriguez, Manuel a. Gómez-Marcos, Emiliano Rodríguez-Sánchez, Jose A Maderuelo-Fernández, Jose A Iglesias-Valiente, Maria C Patino-Alonso, Diana Pérez-Arechaederra, Sara Mora-Simón, Cristina Agudo-Conde, Maria C Castaño-Sánchez, Carmela Rodríguez-Martín, Benigna Sánchez-Salgado, Angela de Cabo-Laso, Rosa Alonso-Domínguez, and Natalia Sánchez-Aguadero.

2. Collaborating centers: Passeig de Sant Joan Health Center (Catalan Health Service): Carlos Martín-Cantera, Jose Canales-Reina, Epifania Rodrigo de Pablo, Maria L Lasaosa-Medina, Maria J Calvo-Aponte, A Rodríguez-Franco, Carmen Martin-Borras, Anna Puig-Ribera, Ruben Colominas-Garrido, and Elisa Puigdomenech Puig.

3. Ca N'Oriac Health Center (Catalan Health Service): Monserrat Romaguera-Bosch.

4. Sant Roc Health Center (Catalan Health Service): Sandra Maneus.

5. Río Tajo Health Center (Health Service of Castilla-La Mancha): Yolanda Schmolling-Guinovart, Beatriz Rodríguez-Martín, Alicia Fernández del Rio, José A Fernández-Díaz, José B Calderón-Ubeda, José L Menéndez-Obregón, Antonio Segura-Fragoso, Carmen Zabala-Baños, Vicente Martínez-Vizcaíno, and María Martínez-Andrés.

6. Casa de Barco Health Center (Health Service of Castilla y León): Maria C Fernández-Alonso, Amparo Gómez-Arranz, Aventina de la Cal de la Fuente, Marta Menéndez-Suarez, Irene Repiso-Gento.

7. San Pablo Health Center (Health Service of Castilla y León): Maria I Arranz-Hernando, Maria I Pérez-Concejo, Maria A Alonso-Manjarres, Maria E Villarroya, Maria J Arribas de Rodrigo, Margarita Pérez de Lis, Maria D de Arriba-Gómez, and Maria M López-Arroyo.

8. Torre Ramona Health Center (Health Service of Aragón): Natividad González-Viejo, Jose F Magdalena-Belio, Luis Otegui-Ilarduya, Francisco J Rubio-Galán, Amor Melguizo-Bejar, Ines Sauras-Yera, Maria J Gil-Train, Marta Iribarne-Ferrer, Olga Magdalena-González, and Miguel A Lafuente-Ripolles.

9. Primary Care Research Unit of Bizkaia (Basque Health Service-Osakidetza): Gonzalo Grandes, Álvaro Sanchez, Verónica Arce, Maria S Arietaleanizbeaskoa, Nere Mendizabal, and Eguskiñe Iturregui-San Nicolas.

10. CGB Computer Company, Salamanca, Spain (contribution to technical development of the app EVIDENT II).

11. Funding/Support: This study was funded by the Spanish Ministry of Science and Innovation (MICINN), the Carlos III Health Institute/European Regional Development Fund (ERDF) (MICINN, ISCIII/FEDER) (FIS: PI13/00618, PI13/01526, PI13/00058, PI13/01635, PI13/02528, PI12/01474; RETICS: RD12/0005, RD16/0007), the Regional Health Management of Castilla y León (GRS 1191/B/15, GRS 909/B/14, GRS 770/B/13), and the Infosalud Foundation. They played no role in the study design, data analysis, reporting of results, or decision to submit the manuscript for publication.

Authors' Contributions

Conception and design: José I Recio-Rodríguez, Jose A Maderuelo-Fernandez, Emiliano Rodriguez-Sanchez, Manuel A Gómez-Marcos, and Luis García-Ortiz. Drafting of the paper: José I Recio-Rodríguez and Luis García-Ortiz. Analysis and interpretation of the data: Luis García-Ortiz and Jose A Maderuelo-Fernandez. Critical revision of the paper for important intellectual content: José I Recio-Rodríguez, Jose A Maderuelo-Fernandez, Emiliano Rodriguez-Sanchez, Manuel A Gómez-Marcos, and Luis García-Ortiz. Final approval of the paper: José I Recio-Rodríguez, Cristina Agudo Conde, Carlos Martín-Cantera, Natividad González-Viejo, María Carmen Fernandez Alonso, Maria S Arietaleanizbeascoa, Yolanda Schmolling-Guinovart, Jose A Maderuelo-Fernandez, Emiliano Rodriguez-Sanchez, Manuel A Gómez-Marcos, and Luis García-Ortiz. Statistical expertise: Luis García-Ortiz and Jose A Maderuelo-Fernandez. Collection and assembly of data: José I Recio-Rodríguez, Cristina Agudo Conde, Carlos Martín-Cantera, Natividad González-Viejo, María Carmen Fernandez Alonso, Maria S Arietaleanizbeascoa, and Yolanda Schmolling-Guinovart.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Baseline adherence to the Mediterranean diet.

[PDF File (Adobe PDF File), 25KB - jmir_v18i12e331_app1.pdf]

Multimedia Appendix 2

Subgroup analyses by age, gender, baseline physical activity and weight.

[PDF File (Adobe PDF File), 32KB - jmir_v18i12e331_app2.pdf]

Multimedia Appendix 3

Subgroup analyses by sociodemographic variables.

[PDF File (Adobe PDF File), 33KB - jmir_v18i12e331_app3.pdf]

Multimedia Appendix 4

7-day PAR in university women subgroup.

[PDF File (Adobe PDF File), 25KB - jmir_v18i12e331_app4.pdf]

References

1. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity--a systematic review of longitudinal studies. BMC Public Health 2013;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
2. Samitz G, Egger M, Zwahlen M. Domains of physical activity and all-cause mortality: systematic review and dose-response meta-analysis of cohort studies. Int J Epidemiol 2011 Oct;40(5):1382-1400 [FREE Full text] [doi: 10.1093/ije/dyr112] [Medline: 22039197]
3. Löllgen H, Böckenhoff A, Knapp G. Physical activity and all-cause mortality: an updated meta-analysis with different intensity categories. Int J Sports Med 2009 Mar;30(3):213-224. [doi: 10.1055/s-0028-1128150] [Medline: 19199202]

4. Kodama S, Tanaka S, Heianza Y, Fujihara K, Horikawa C, Shimano H, et al. Association between physical activity and risk of all-cause mortality and cardiovascular disease in patients with diabetes: a meta-analysis. Diabetes Care 2013 Feb;36(2):471-479 [FREE Full text] [doi: 10.2337/dc12-0783] [Medline: 23349151]
5. Rossi A, Dikareva A, Bacon SL, Daskalopoulou SS. The impact of physical activity on mortality in patients with high blood pressure: a systematic review. J Hypertens 2012 Jul;30(7):1277-1288. [doi: 10.1097/HJH.0b013e3283544669] [Medline: 22573122]
6. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
7. Tucker JM, Welk GJ, Beyler NK. Physical activity in US: adults compliance with the Physical Activity Guidelines for Americans. Am J Prev Med 2011 Apr;40(4):454-461. [doi: 10.1016/j.amepre.2010.12.016] [Medline: 21406280]
8. García-Ortiz L, Recio-Rodríguez JI, Schmidt-Trucksäss A, Puigdomenech-Puig E, Martínez-Vizcaíno V, Fernández-Alonso C, et al. Relationship between objectively measured physical activity and cardiovascular aging in the general population--the EVIDENT trial. Atherosclerosis 2014 Apr;233(2):434-440. [doi: 10.1016/j.atherosclerosis.2014.01.021] [Medline: 24530775]
9. Sanchez A, Bully P, Martinez C, Grandes G. Effectiveness of physical activity promotion interventions in primary care: a review of reviews. Prev Med 2015 Jul;76 Suppl:S56-S67. [doi: 10.1016/j.ypmed.2014.09.012] [Medline: 25263343]
10. Estruch R, Ros E, Salas-Salvadó J, Covas M, Corella D, Arós F, et al. Primary prevention of cardiovascular disease with a Mediterranean diet. N Engl J Med 2013 Apr 4;368(14):1279-1290. [doi: 10.1056/NEJMoa1200303] [Medline: 23432189]
11. Schwingshackl L, Missbach B, König J, Hoffmann G. Adherence to a Mediterranean diet and risk of diabetes: a systematic review and meta-analysis. Public Health Nutr 2015 May;18(7):1292-1299. [doi: 10.1017/S1368980014001542] [Medline: 25145972]
12. Patino-Alonso MC, Recio-Rodríguez JI, Belio-Magdalena JF, Colominas-Garrido R, Lema-Bartolomé J, Arranz-Gomez A, et al. Factors associated with adherence to the Mediterranean diet in the adult population. J Acad Nutr Diet 2014 Apr;114(4):583-589. [doi: 10.1016/j.jand.2013.07.038] [Medline: 24209889]
13. Maderuelo-Fernandez JA, Recio-Rodríguez JI, Patino-Alonso MC, Pérez-Arechaederra D, Rodriguez-Sanchez E, Gomez-Marcos MA, et al. Effectiveness of interventions applicable to primary health care settings to promote Mediterranean diet or healthy eating adherence in adults: A systematic review. Prev Med 2015 Jul;76 Suppl:S39-S55. [doi: 10.1016/j.ypmed.2014.12.011] [Medline: 25524613]
14. Kamel-Boulos MN, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. Biomed Eng Online 2011;10:24 [FREE Full text] [doi: 10.1186/1475-925X-10-24] [Medline: 21466669]
15. Higgins JP. Smartphone applications for patients' health and fitness. Am J Med 2016 Jan;129(1):11-19. [doi: 10.1016/j.amjmed.2015.05.038] [Medline: 26091764]
16. Flores-Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
17. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-S12. [doi: 10.7326/M13-3005] [Medline: 25402403]
18. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
19. Modave F, Bian J, Leavitt T, Bromwell J, Harris IC, Vincent H. Low Quality of Free Coaching Apps With Respect to the American College of Sports Medicine Guidelines: A Review of Current Mobile Apps. JMIR Mhealth Uhealth 2015;3(3):e77 [FREE Full text] [doi: 10.2196/mhealth.4669] [Medline: 26209109]
20. Recio-Rodríguez JI, Martín-Cantera C, González-Viejo N, Gómez-Arranz A, Arietaleanizbeascoa MS, Schmolling-Guinovart Y, et al. Effectiveness of a smartphone application for improving healthy lifestyles, a randomized clinical trial (EVIDENT II): study protocol. BMC Public Health 2014;14:254 [FREE Full text] [doi: 10.1186/1471-2458-14-254] [Medline: 24628961]
21. García-Ortiz L, Recio-Rodríguez JI, Martín-Cantera C, Cabrejas-Sánchez A, Gómez-Arranz A, González-Viejo N, et al. Physical exercise, fitness and dietary pattern and their relationship with circadian blood pressure pattern, augmentation index and endothelial dysfunction biological markers: EVIDENT study protocol. BMC Public Health 2010;10:233 [FREE Full text] [doi: 10.1186/1471-2458-10-233] [Medline: 20459634]
22. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
23. Grandes G, Sanchez A, Ortega Sanchez-Pinilla R, Torcal J, Montoya I, Lizarraga K, et al. Effectiveness of physical activity advice and prescription by physicians in routine primary care: a cluster randomized trial. Arch Intern Med 2009 Apr 13;169(7):694-701. [doi: 10.1001/archinternmed.2009.23] [Medline: 19364999]

24. Matthews CE, Keadle SK, Sampson J, Lyden K, Bowles HR, Moore SC, et al. Validation of a previous-day recall measure of active and sedentary behaviors. Med Sci Sports Exerc 2013 Aug;45(8):1629-1638 [FREE Full text] [doi: 10.1249/MSS.0b013e3182897690] [Medline: 23863547]
25. Melanson EL, Freedson PS. Validity of the Computer Science and Applications, Inc (CSA) activity monitor. Med Sci Sports Exerc 1995 Jun;27(6):934-940. [Medline: 7658958]
26. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity (Silver Spring) 2007 Oct;15(10):2371-2379 [FREE Full text] [doi: 10.1038/oby.2007.281] [Medline: 17925461]
27. Chen KY, Bassett DR. The technology of accelerometry-based activity monitors: current and future. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S490-S500. [Medline: 16294112]
28. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [Medline: 9588623]
29. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1423-1434. [doi: 10.1249/mss.0b013e3180616b27] [Medline: 17762377]
30. Schröder H, Fitó M, Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, et al. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr 2011 Jun;141(6):1140-1145 [FREE Full text] [doi: 10.3945/jn.110.135566] [Medline: 21508208]
31. Partridge SR, McGeechan K, Hebden L, Balestracci K, Wong AT, Denney-Wilson E, et al. Effectiveness of a mHealth lifestyle program with telephone support (TXT2BFiT) to prevent unhealthy weight gain in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2015;3(2):e66 [FREE Full text] [doi: 10.2196/mhealth.4530] [Medline: 26076688]
32. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014 Jul;64(624):e384-e391. [doi: 10.3399/bjgp14X680461] [Medline: 24982490]
33. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Hardy D, Smith SA. Smartphone applications for promoting healthy diet and nutrition: a literature review. Jacobs J Food Nutr 2015;2(3):021 [FREE Full text] [Medline: 26819969]
34. Wang Q, Egelandsdal B, Amdam G, Almli V, Oostindjer M. Diet and physical activity apps: perceived effectiveness by app users. JMIR Mhealth Uhealth 2016;4(2):e33. [doi: 10.2196/mhealth.5114] [Medline: 27056639]

Abbreviations

7-day PAR: 7-day Physical Activity Recall
MEDAS: Mediterranean Diet Adherence Screener
MET: metabolic equivalent
MVPA: moderate-to-vigorous physical activity
PA: physical activity

Edited by G Eysenbach; submitted 15.10.16; peer-reviewed by M Harris, A Wong; comments to author 03.11.16; revised version received 18.11.16; accepted 23.11.16; published 19.12.16.

Please cite as:
Recio-Rodriguez JI, Agudo-Conde C, Martin-Cantera C, González-Viejo MN, Fernandez-Alonso MDC, Arietaleanizbeaskoa MS, Schmolling-Guinovart Y, Maderuelo-Fernandez JA, Rodriguez-Sanchez E, Gomez-Marcos MA, Garcia-Ortiz L, EVIDENT Investigators
Short-Term Effectiveness of a Mobile Phone App for Increasing Physical Activity and Adherence to the Mediterranean Diet in Primary Care: A Randomized Controlled Trial (EVIDENT II Study)
J Med Internet Res 2016;18(12):e331
URL: http://www.jmir.org/2016/12/e331/
doi:10.2196/jmir.6814
PMID:27993759

Original Paper

The Use and Effectiveness of Mobile Apps for Depression: Results From a Fully Remote Clinical Trial

Patricia A Arean[1], PhD; Kevin A Hallgren[1], PhD; Joshua T Jordan[2,3], MA; Adam Gazzaley[2,3], MD, PhD; David C Atkins[4], PhD; Patrick J Heagerty[4], PhD; Joaquin A Anguera[2,3], PhD

[1]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States
[2]Department of Neurology, University of California San Francisco, San Francisco, CA, United States
[3]Department of Psychiatry, University of California San Francisco, San Francisco, CA, United States
[4]Department of Biostatistics in the School of Public Health, University of Washington, Seattle, WA, United States

Corresponding Author:
Joaquin A Anguera, PhD
Department of Neurology
University of California San Francisco
675 Nelson Rising Lane
San Francisco, CA, 94158
United States
Phone: 1 415 502 7322
Email: joaquin.anguera@ucsf.edu

Abstract

Background: Mobile apps for mental health have the potential to overcome access barriers to mental health care, but there is little information on whether patients use the interventions as intended and the impact they have on mental health outcomes.

Objective: The objective of our study was to document and compare use patterns and clinical outcomes across the United States between 3 different self-guided mobile apps for depression.

Methods: Participants were recruited through Web-based advertisements and social media and were randomly assigned to 1 of 3 mood apps. Treatment and assessment were conducted remotely on each participant's smartphone or tablet with minimal contact with study staff. We enrolled 626 English-speaking adults (≥18 years old) with mild to moderate depression as determined by a 9-item Patient Health Questionnaire (PHQ-9) score ≥5, or if their score on item 10 was ≥2. The apps were (1) Project: EVO, a cognitive training app theorized to mitigate depressive symptoms by improving cognitive control, (2) iPST, an app based on an evidence-based psychotherapy for depression, and (3) Health Tips, a treatment control. Outcomes were scores on the PHQ-9 and the Sheehan Disability Scale. Adherence to treatment was measured as number of times participants opened and used the apps as instructed.

Results: We randomly assigned 211 participants to iPST, 209 to Project: EVO, and 206 to Health Tips. Among the participants, 77.0% (482/626) had a PHQ-9 score >10 (moderately depressed). Among the participants using the 2 active apps, 57.9% (243/420) did not download their assigned intervention app but did not differ demographically from those who did. Differential treatment effects were present in participants with baseline PHQ-9 score >10, with the cognitive training and problem-solving apps resulting in greater effects on mood than the information control app ($\chi 22=6.46$, *P*=.04).

Conclusions: Mobile apps for depression appear to have their greatest impact on people with more moderate levels of depression. In particular, an app that is designed to engage cognitive correlates of depression had the strongest effect on depressed mood in this sample. This study suggests that mobile apps reach many people and are useful for more moderate levels of depression.

ClinicalTrial: Clinicaltrials.gov NCT00540865; https://www.clinicaltrials.gov/ct2/show/NCT00540865 (Archived by WebCite at http://www.webcitation.org/6mj8IPqQr)

(J Med Internet Res 2016;18(12):e330) doi:10.2196/jmir.6482

KEYWORDS

depression; mobile apps; RCT; randomized controlled trial; cognitive training; iPST; problem-solving therapy

Introduction

Major depressive disorder affects approximately 7% of the US population in a given year [1], and approximately 16% of all individuals will experience major depression at least once in their lifetime [2]. Despite the fact that depression is treatable [3,4], less than a quarter of individuals receive adequate care for this illness [1] due to treatment access barriers, such as time and transportation constraints, long waiting lists, and a dearth of trained professionals to provide high-quality care [5-7]. Access problems could be easily mitigated through the use of technology; several studies have already demonstrated that telemedicine and Internet-based approaches are feasible and as effective as in-person treatment [8,9]. The success of these distance approaches has resulted in considerable interest in the use of mobile phone apps as an alternative care delivery platform. Not only do mental health apps have tremendous reach, but also patients can access these tools whenever they feel the need and as often as they like without having to wait until a mental health professional is available [10,11]. Such reach is evident when considering that 68% of all adults in the United States own a smartphone, and approximately 45% own a tablet device [12].

Hundreds of apps for depression are available for download on one’s smart device [13-15], with the majority of these apps designed to be self-guided. While a few proof-of-concept studies and small-scale randomized controlled trials (RCTs) have been conducted, to our knowledge, none have compared theoretically driven interventions versus controls, nor have they investigated the effects of these apps under real-world conditions [16]. Furthermore, we know little about the people who download depression treatment apps in terms of symptom severity, disability, and use of more traditional mental health services. Although a few proof-of-concept studies have found that adherence to depression app guidelines tends to be poor over time [17-21], there is little information about the relative use patterns between different types of apps. We conducted a large-scale RCT (NCT00540865) of 3 different depression apps to answer the following questions: (1) Who downloads mobile apps for depression? (2) How do people who download these apps use them? (3) What is the impact of these apps as they are typically used? (4) What are the methodological issues inherent in conducting a fully remote RCT? We have already published the data on the methodological issues in the recruitment and retention of participants into a large-scale, remote RCT [10]. We report here on the characteristics of people who download and use depression apps, what their use patterns are like, and how effective these interventions are in light of typical use patterns typically deployed in the community.

Methods

Study Design

This was a fully remote, randomized clinical field trial comparing 2 active apps and a control app for mood [10], and to test the feasibility of remote research using mobile phone apps. We recruited participants through a variety of Web-based advertising sources, including Craigslist (Craigslist Inc, San Francisco, CA, USA), Google AdWords (Google Inc, Mountain View, CA, USA), and social media outlets (eg, Twitter; Twitter, Inc, San Francisco, CA, USA) from all 50 US states. All potential participants learned about the study through a website explaining study details, which led to a screening protocol using automated software (SurveyGizmo; Widgix, LLC, Boulder, CO, USA) to determine eligibility. Informed consent was conducted through a 2-minute video explaining the study risks and benefits, and the payment structure over the course of the study. In addition to a PDF file of the consent form was provided. Participants had to complete a 3-item quiz testing their understanding of the study to advance to the randomization phase.

Participant contact was minimal. Study staff contacted participants via email or short message service (SMS) text messaging to remind them to use their intervention or assessment app if they had 3 consecutive days of missing data. Aside from this, participants were contacted only when they (1) were due with a payment, or (2) reached out to study staff for technical support. Contact was primarily through email or SMS, with occasional phone call meetings if needed to help participants download their intervention apps. Ethical approval for the trial was granted by the University of California, San Francisco, Committee for Human Research.

Participants

To be eligible, participants had to be English speakers, be at least 18 years old, and own a smartphone (iPhone or an Android device) with Wi-Fi or third- or fourth-generation capabilities. Because 1 of the 2 interventions was available only on the iOS mobile operating system and had to be used on devices with a visual field larger than that on a typical smartphone, participants had to own an Apple iPad 2.0 or newer version.

Participants had to obtain a score of 5 or more on the 9-item Patient Health Questionnaire (PHQ-9 [22]) or a score of 2 or greater on item 10 of the PHQ-9 (“If you checked off any problems, how difficult have these problems made it for you to do your work, take care of things at home, or get along with other people?”). The decision to include participants with mild symptoms of depression, rather than limit the sample to those with moderately severe depression (eg, PHQ-9 score >10), was based on our intent to test the use and effects of these apps in people with a range symptoms of depression, as well as to determine what proportion of participants who downloaded depression apps fell into mild, moderate, or severe ranges of depression. The only exception to this rule was suicidal ideation. Participants with a PHQ-9 suicide item score of 1 or more were referred to the suicide hotline. We randomly assigned participants to 1 of the 3 apps using a random number generator built into the eligibility survey.

Baseline Assessments

We collected information on demographics, depression severity using the PHQ-9 [22], functional disability using the Sheehan Disability Scale (SDS [23]), anxiety using the Generalized Anxiety Disorder 7-item scale (GAD-7 [24]), history of mania or psychosis using the Improving Mood-Promoting Access to Collaborative Treatment (IMPACT) assessment of mania and psychosis [25], and alcohol use using the Alcohol Use Disorders Identification Test (AUDIT-C) [26]. We also collected data on self-reported quality of sleep, current use of mobile phone apps, and engagement in outside mental health treatment. Participants in all groups were sent these surveys and were paid US $15.00 for completing the baseline assessment. All treatments and assessments were delivered over the participants' smart devices. Human interaction was limited to reminders sent via SMS or email based on each participant's stated preferences.

Procedures for App Access

Once participants completed the consent process, a secure, 1-user valid link to a secure webpage was sent to participants' email addresses that contained a brief personalized YouTube video explaining how to download and then use their assigned intervention. This webpage also contained a link to automatically download said apps to the participants' phone or iPad.

Intervention Apps

Cognitive Control App (Project: EVO)

Participants randomly assigned to this condition were encouraged to use Project: EVO (Akili Interactive Labs, Larkspur, CA, USA) 6 times a week for approximately 30 minutes per day. This app is designed as a video game that modulates cognitive control abilities, a common neurological deficit seen in depression [27]. The app uses adaptive algorithms to adjust the intervention's difficulty to the user's level of proficiency over time. Previous work that Project: EVO was derived from demonstrated at this dosage that its prescribed use could improve cognitive control in older adults [28], with preliminary evidence for a similar effect on depression [29].

Problem-Solving Therapy App (iPST)

Participants were to use the problem-solving app iPST as often as possible each week, with a minimum of once per week being the typical amount undertaken in a clinical setting. iPST is based on problem-solving therapy, which focuses on a 7-step model to manage mood. In this app, participants choose a goal and are guided through a 7-step process to create an action plan. Problem-solving therapy is an evidence-based treatment [30] and is particularly effective in treating depression [3,31].

Information Control (Health Tips)

Participants in this condition were given an app that provided daily health tips for improved health, such as self-care (eg, taking a shower) or physical activity (eg, taking a walk). Although it provided daily advice on improving one's health, it is not tied to any specific theory, similar to supportive control treatments. Participants were not required to act on the health tip.

App Design

All 3 apps were developed by professionals with user-centered design method experience (Project: EVO, Akili Interactive Labs, Larkspur, CA, USA; iPST, Wow Internet Labz Pvt Ltd, Bengaluru, India; and Health Tips, Ginger.io, San Francisco, CA, USA) to maximize engagement and minimize app use burden, a common problem associated with app adherence [32,33]. Participants were expected to use their assigned app as instructed for 1 month. The survey app had internally programmed reminders to notify the user that a new assessment was ready for completion, or that they had not completed a given assessment 8 hours after it was originally transmitted. For each intervention app, our team of research assistants monitored participants' use of their assigned apps using a custom Web-based dashboard. If a participant had not used Project: EVO or iPST in 72 hours, they were sent an email or SMS (based on their indicated preference) reminding them to use their assigned app. If participants did not use their app within the next 72 hours, no further reminders were sent. App use was collected and ported to a secure data server at the University of California, San Francisco, which met all Health Insurance Portability and Accountability Act (HIPAA) and security requirements imposed by the university. Participants were not compensated for using the apps. Note that the eligibility criteria for randomization required that participants had (1) either an iPhone or Android smartphone, and (2) an iPad 2.0 or newer. Participants who met criteria as per requirement (1) but not (2) were given Project: EVO if they had an iPhone, or Health Tips if they had an Android.

Outcome Assessments

The primary outcome measures were of depression (PHQ-9) and function (SDS [34]), with these scores captured weekly for the first 4 weeks of treatment, then at 8 and 12 weeks (see Multimedia Appendix 1 for discussion of other exploratory outcomes). Participants were paid US $20.00 for completing assessments at the 4-, 8-, and 12-week marks. Because all assessment was conducted using assessment software, procedures for blinding research assistants was not necessary.

Data Analysis

All analyses were modeled on an intent-to-treat approach. We used hurdle models [35] to estimate (1) predictors of app adherence and follow-up rates, (2) which variables would predict using the apps at least once (as odds ratios, ORs) and, (3) which variables would predict the number of times the apps were used (as rate ratios, RRs). ORs and RRs used standardized z scores as continuous predictors to facilitate interpretations associated with the relative increase in odds or rate associated with a 1-SD change on given predictors. Predictors in hurdle models included baseline PHQ-9, SDS, GAD-7, AUDIT-C, IMPACT, age, sex, level of education, marital status, employment status, minority status, whether participants endorsed other psychiatric or psychotherapeutic treatments at baseline, and condition assignment.

We characterized participant adherence in 3 ways: none=no use at all (downloading the app did not count toward use); suboptimal=some use, but never met adherence criteria for a given week *or* only met adherence criteria for 1 of the 4 weeks; and optimal=met adherence criteria for at least 2 of the 4 weeks, such that participants exactly followed the specified instructions for app use as outlined in the informative YouTube videos. For Project: EVO, we considered three 30-minute sessions per week (or 50% of the indicated amount) to be the minimum acceptable amount of treatment with respect to assessing the feasibility of self-administering a daily cognitive intervention. We used identical procedures to predict follow-up rates, but included participants from all 3 treatment conditions.

To estimate changes in depression and disability during and after the treatment period, we tested growth curve models using multilevel modeling with continuous piecewise growth curves for each period. These models used restricted maximum likelihood with all available data to reduce missing data bias [36], and included random intercepts and random effects for time. Growth curve models are well known to be better than some other methods for estimating interindividual variability in intraindividual patterns of change, including accounting for missing data in a rigorous manner [37]. We entered baseline PHQ-9 or SDS scores as control variables along with medication use due to differences between groups in this variable at baseline. Separate models tested the impact of the app conditions on remission rates. Remission was assessed by characterizing the proportion of participants who demonstrated a reduction of 50% of their pretreatment depression and disability scores [38,39].

We conducted sensitivity analyses to evaluate whether missing data biased estimates through pattern-mixture models [40], which is considered the gold standard for RCT studies [41,42]. This approach is deemed important to test the assumption that data are missing at random for multiple imputation methods [43]. We also tested whether the amount of change in depression or disability was moderated by baseline PHQ-9, GAD-7, SDS, AUDIT-C, age, or app use through baseline variable-by-time interactions.

Sample Size and Power

A power analysis [44] indicated that 200 participants per condition would provide 0.80 power to detect whether an active treatment condition improved by 2 points on the PHQ-9 beyond the control condition (approximately Cohen *d*=0.4), assuming a 50% dropout and an alpha level of .05.

Results

Participant Flow, Recruitment, and Baseline Data

National recruitment began in August 2014, with 2923 participants screened across the 5 waves of 2-week advertising (total of 5 months of recruitment; see Consolidated Standards of Reporting Trials [CONSORT] diagram, Figure 1). A total of 626 participants had both an iPad 2.0 and a smartphone, and were randomly assigned to the 3 study arms (iPST, 211; Project: EVO, 209; Health Tips, 206). The mean age of the sample was 33.95 (SD 11.84) years, and the mean PHQ-9 score at baseline (13.64, SD 4.95) indicated the sample was moderately depressed. The proportion of individuals in our sample with a PHQ-9 total score between 5 and 10 was 23.0% (144/626), while the proportion of those with a score ≥10 was 77.0% (482/626). Only 11 participants who had a total PHQ-9 score less than 5 reported a score greater than 2 on item 10 of this tool, with their baseline total PHQ-9 being 3.09 (SD 0.83). The majority of the sample was female (494/626, 79.0%) and non-Hispanic white (376/626, 60.1%) Table 1 presents the demographic characteristics of the randomized sample, including ethnic group proportions, concurrent clinical diagnoses, and those in treatment.

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) diagram. EVO: Project: EVO; HT: Health Tips; PHQ-9: 9-item Patient Health Questionnaire; wk: week.

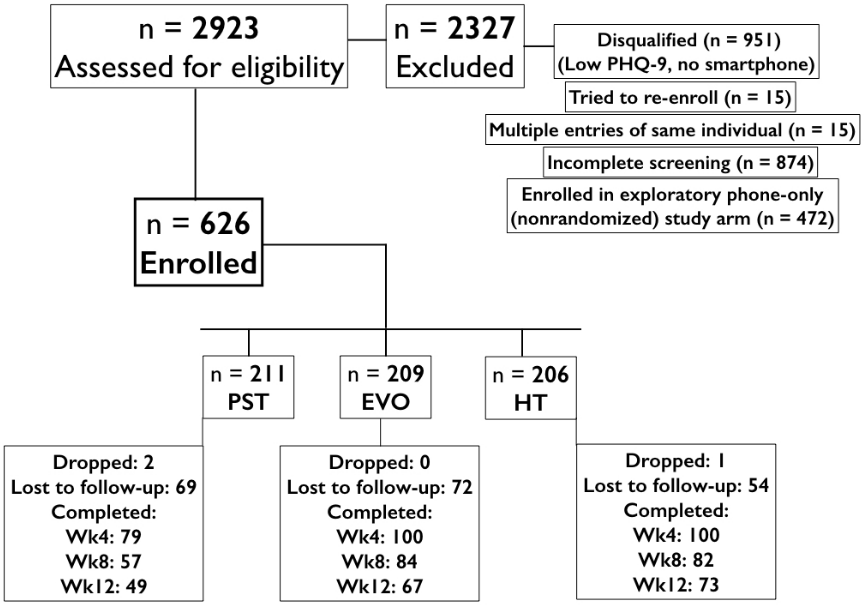


Table 1. Sample descriptive statistics of participants using 3 different self-guided mobile apps for depression.

	EVO[a] (n=209)		iPST[b] (n=211)		HT[c] (n=206)		Total (n=626)	
Baseline variable								
PHQ-9[d] score, mean (SD)	13.76	(4.9)	13.51	(5.1)	13.64	(4.9)	13.64	(4.95)
PHQ-disability score, mean (SD)	1.34	(0.75)	1.44	(0.68)	1.40	(0.73)	1.39	(0.72)
Age in years, mean (SD)	34.9	(12.3)	33.4	(10.9)	33.6	(12.3)	33.9	(11.84)
Male, n (%)	51	(24.2)	48	(23.0)	33	(16.0)	132	(21.1)
University education, n (%)	128	(60.7)	140	(67.0)	130	(63.1)	398	(63.6)
Married, n (%)	62	(29.4)	73	(34.9)	68	(33.0)	203	(32.4)
Employed, n (%)	144	(68.2)	156	(74.6)	141	(68.4)	441	(70.5)
Racial/ethnic minority, n (%)	128	(60.7)	122	(58.4)	124	(60.2)	374	(59.7)
African American	29	(13.7)	29	(13.9)	28	(13.6)	86	(13.7)
American Indian	4	(1.9)	1	(0.5)	1	(0)	6	(1.0)
Asian	22	(10.4)	16	(7.7)	16	(7.8)	54	(8.6)
White	137	(64.9)	140	(67.0)	133	(64.6)	410	(65.5)
>1 race	19	(9.0)	21	(10.0)	26	(12.6)	66	(10.5)
Native Hawaiian/Pacific Islander	0	(0)	2	(1.0)	2	(1.0)	4	(0.6)
Hispanic (any race)	24	(11.4)	33	(15.8)	22	(10.6)	79	(12.6)
First-use variable								
SDS[e], mean (SD)	15.9	(7.1)	14.9	(6.7)	15.9	(7.1)	15.6	(6.96)
GAD-7[f], mean (SD)	10.4	(4.9)	9.2	(4.9)	10.4	(5.3)	10.0	(5.09)
AUDIT-C[g], mean (SD)	3.20	(2.5)	3.03	(2.2)	3.40	(2.3)	3.21	(2.38)
Psychotic symptoms, n (%)	29	(19)	24	(20)	32	(24)	155	(21)
Other psychiatric treatment, n (%)	84	(58)	64	(52)	74	(56)	222	(56)
Psychiatrist	41	(19.6)	31	(14.8)	35	(16.7)	107	(17.1)
Therapist	41	(19.6)	32	(15.3)	30	(14.4)	103	(16.5)
Group	15	(7.1)	10	(4.8)	5	(2.4)	30	(4.8)
Book	26	(12.4)	23	(11.0)	35	(16.7)	84	(13.4)
Medication[h]	63	(30.1)	43	(20.6)	39	(18.6)	145	(23.2)

[a] EVO: Project: EVO.

[b] iPST: problem-solving therapy app.

[c] HT: Health Tips.

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

[e] SDS: Sheehan Disability Scale.

[f] GAD-7: Generalized Anxiety Disorder 7-item scale.

[g] AUDIT-C: Alcohol Use Disorders Identification Test.

[h] EVO had a significantly higher rate of baseline medication use than iPST and HT.

Sensitivity Analysis and Moderators of Treatment Effects

The sensitivity analyses performed here to account for potential bias in the data revealed that the patterns of missing data did not predict significant differences in PHQ-9 (χ^2_6=8.47, *P*=.20) or SDS trajectories (χ^2_6=9.67, *P*=.14) during weeks 1-4. This suggests that overall levels and changes in depression during weeks 1-4 were not significantly different between those who did and did not provide follow-up data after week 4. Comparing participants who used the Project: EVO and iPST apps optimally, suboptimally, or not at all, we observed that the overall levels of depression as scored by the PHQ-9 were not different for the suboptimal or optimal groups relative to the "none" group for both Project: EVO (all *P* ≥.22) and iPST (all

P ≥.64; see Multimedia Appendix 1). Baseline functioning, alcohol use, and age did not significantly moderate the effects of treatment condition on depression outcomes. We found that baseline depression significantly moderated changes in depression over time during weeks 4-12 for the iPST group relative to control (*P*=.02). Baseline anxiety significantly moderated changes in the SDS scores over time during weeks 4-12 for iPST (*P*=.01) but not for Project: EVO (*P*=.08) relative to control.

Assessment Adherence

Of the participants who self-reported a PHQ-9 total score <10, 66.0% (95/144) used the survey app, compared with 64.9% (313/482) of those with a PHQ-9 score >10, with this difference in use being nonsignificant (*P*=.82). A total of 354/626 (56.6%) provided at least one follow-up assessment. Participants with at least one follow-up assessment provided an average of 5.83 (SD 2.42) of 8 possible follow-up assessments, with participants who were older (RR 1.08, *P*=.02) completing a greater number of follow-up assessments (see Table 2). Racial/ethnic minorities were more likely than nonminorities to provide at least one follow-up assessment (OR 1.26, *P*<.001; see Multimedia Appendix 1 for more on follow-up analyses, participant expectancy, and perceived study burden).

Table 2. Predictors of app use and follow-up completion.

	Predictors of app use					Predictors of follow-up				
	Any use		Use count		*P* value	Any follow-up		No. of follow-ups		*P* value
	OR[a]	(SE[b])	RR[c]	(SE)		OR	(SE)	Beta	(SE)	
Baseline variable										
PHQ-9[d]	1.01	(0.12)	0.79	(0.09)	.93	0.90	(0.09)	–.23	(0.16)	.04
PHQ-disability	0.93	(0.11)	1.27	(0.15)	.54	0.96	(0.09)	–.17	(0.16)	.03
Age	1.06	(0.12)	1.15	(0.12)	.61	1.08	(0.10)	.29	(0.15)	.19
Male	0.89	(0.09)	0.96	(0.09)	.25	0.96	(0.08)	–.11	(0.14)	.67
University education	1.21	(0.13)	1.01	(0.11)	.07	1.01	(0.09)	.18	(0.14)	.92
Married	0.93	(0.10)	0.93	(0.09)	.53	0.84	(0.07)	–.21	(0.14)	.46
Employed	1.09	(0.11)	0.92	(0.09)	.42	0.97	(0.08)	–.05	(0.14)	.39
Racial/ethnic minority	1.04	(0.11)	0.95	(0.09)	.70	1.26	(0.11)	.25	(0.14)	.57
Condition: iPST[e]	1.06	(0.21)	1.11	(0.21)	.78	1.05	(0.21)	–.05	(0.33)	.60
Condition: EVO[f]						0.69	(0.14)	–.93	(0.33)	
First-use variable										
SDS[g]	0.97	(0.16)	1.30	(0.16)	.83	1.29	(0.24)	–.06	(0.19)	.04
GAD-7[h]	1.03	(0.17)	0.78	(0.10)	.86	0.76	(0.14)	–.26	(0.19)	.04
AUDIT-C[i]	1.03	(0.13)	0.77	(0.06)	.80	0.96	(0.14)	–.14	(0.15)	.001
Psychotic symptoms	1.19	(0.40)	1.02	(0.25)	.62	1.34	(0.53)	.24	(0.39)	.93
Other psychiatric treatment	0.77	(0.21)	1.08	(0.21)	.34	0.73	(0.23)	–.03	(0.32)	.71

[a] OR: odds ratio.

[b] SE: standard error.

[c] RR: rate ratio.

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

[e] iPST: problem-solving therapy app.

[f] EVO: Project: EVO.

[g] SDS: Sheehan Disability Scale.

[h] GAD-7: Generalized Anxiety Disorder 7-item scale.

[i] AUDIT-C: Alcohol Use Disorders Identification Test.

Intervention Adherence

Because the Health Tips control did not require interaction, we report adherence for the active apps only. Among the 420 participants in the Project: EVO and iPST conditions, 243 (57.9%) did not download their assign app. Those who used their app at least once used it on average 10.78 (SD 11.44) times. Higher baseline depression (PHQ-9) and anxiety (GAD-7) were associated with less use, such that a 1-SD increase in each was associated with a 21% and 23% lower rate of adherence, respectively (RR 0.79 and 0.78, respectively, *P*=.04 and *P*=.001, respectively). However, higher disability was associated with a greater adherence (27% for PHQ-disability scale and 30% for SDS; RR 1.27 and 1.30, respectively, *P*=.047 and *P*=.04; see Table 2). None of the baseline variables were significantly associated with the likelihood of using versus not using the Project: EVO or iPST apps. Among participants with at least one use, higher baseline PHQ-9, GAD-7, and AUDIT-C scores were associated with fewer uses, while higher PHQ-disability scores and SDS scores were associated with a greater number of uses.

We further tested whether there were condition-by-baseline variable interactions that predicted use counts and the likelihood of at least one use. Marital status also interacted with treatment condition to predict having at least one instance of use (*P*=.02), such that married individuals were less likely to use iPST once (OR 0.54, *P*=.05). Baseline AUDIT-C scores significantly interacted with treatment condition to predict use counts among those with at least one use (*P*=.048). Specifically, higher AUDIT-C scores were significantly associated with lower use counts in the Project: EVO condition (RR 0.59, *P*=.001). There were no significant interactions between treatment and the variables presented in Table 2 predicting use.

Baseline depression significantly interacted with treatment condition to predict the likelihood of having at least one use instance (*P*=.01), such that higher baseline depression was associated with a lower likelihood of using the Project: EVO app at least once (OR 0.73, *P*=.03; 95% CI 0.55-0.96). All app adherence significantly declined over time (log OR –0.77, SE 0.23, *z* score –3.30, *P*<.001), with no main effect of group (log OR 0.80, SE 0.51, *z* score 1.56, *P*=.12) or group-by-time interaction (log OR 0.24, SE 0.31, *z* score 0.77, *P*=.44; see Figure 2, parts a and b).

Figure 2. Participant intervention use and changes in primary outcome measures over time. (a) Average number of active intervention uses across the first 4 weeks of the study. (b) Number of participants using each active intervention by the level of adherence. (c) 9-item Patient Health Questionnaire (PHQ-9) depression scores over time for each intervention. (d) Sheehan Disability Scale scores over time for each intervention. Control: Health Tips; EVO: Project: EVO; iPST: problem-solving therapy app. Error bars indicate +/- 1 SE.

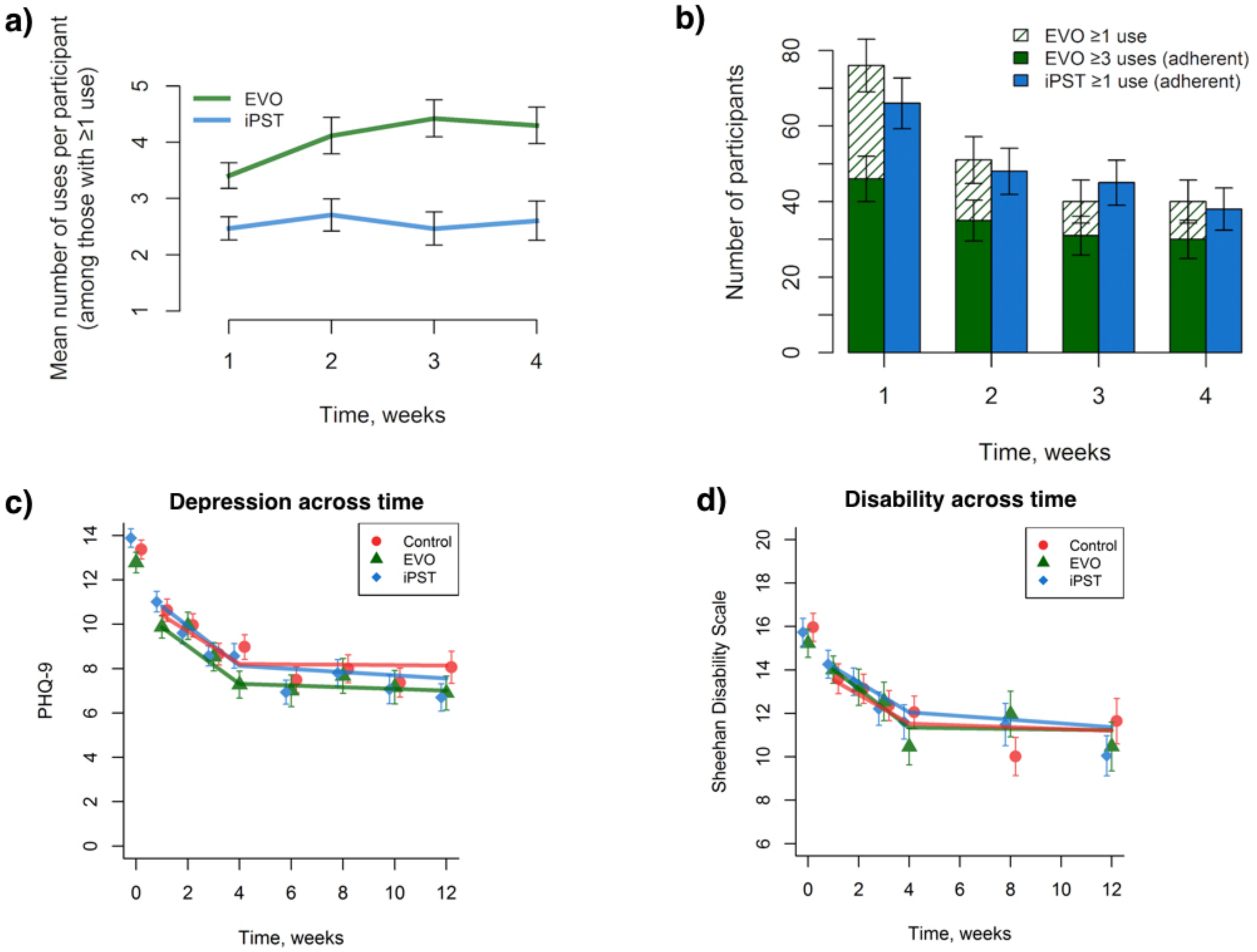


Table 3. Main effects of treatment on changes in depression and disability.

Fixed effects	PHQ-9[a]			SDS[b]		
	B	SE[c]	*P* value	B	SE	*P* value
(Intercept, week 4)	8.21	(0.46)	.00	11.33	(0.64)	.000
Baseline PHQ-9 or SDS	2.89	(0.22)	.00	3.91	(0.26)	.000
Medication use	1.10	(0.43)	.01	0.66	(0.53)	.21
Change per week, weeks 1-4	–0.73	(0.14)	.00	–0.67	(0.22)	.002
EVO[d] condition	–0.52	(0.67)	.44	0.05	(0.94)	.96
iPST[e] condition	–0.60	(0.62)	.34	0.47	(0.89)	.60
Change per week, weeks 4-12	–0.01	(0.06)	.90	–0.04	(0.10)	.69
EVO × change per week, weeks 1-4	–0.12	(0.21)	.58	–0.22	(0.33)	.50
iPST × change per week, weeks 1-4	–0.16	(0.19)	.40	–0.02	(0.31)	.94
EVO × change per week, weeks 4-12	–0.03	(0.10)	.77	0.03	(0.16)	.87
iPST × change per week, weeks 4-12	–0.06	(0.09)	.49	–0.05	(0.15)	.76

[a] PHQ-9: 9-item Patient Health Questionnaire.

[b] SDS: Sheehan Disability Scale.

[c] SE: standard error.

[d] EVO: Project: EVO.

[e] iPST: problem-solving therapy app.

Depression Outcomes

Depression Symptom Severity

For the total sample, PHQ-9 scores decreased an average of 0.73 points per week during the treatment phase and did not significantly change from weeks 4-12 (see Figure 2 c). The models revealed no significant differences between Project: EVO and iPST compared with control at week 4 or week 12, and did not differ in rates of change over time (see Table 3). With respect to treatment remission (using a reduction of pretreatment scores of at least 50% as the criterion for this), 45.0% (45/100) of Project: EVO participants and 46% (36/79) of iPST participants showed improvement by the 4-week assessment, as compared with 34.0% (34/100) of the control app (χ^2=3.36, *P*=.19).

Outcomes by Baseline Depression Severity

A total of 144 participants were classified as having mild symptoms of depression (PHQ-9 ≤9) and 482 with moderate symptoms of depression at baseline (PHQ-9 ≥10). For the mildly depressed subgroup, Project: EVO and iPST did not significantly differ from the control condition at any point. For the subgroup with higher baseline depression, depression was significantly lower at week 12 for the iPST condition (difference=1.79, SE 0.76, t_{201}=–2.36, *P*=.02) but not the Project: EVO condition (*P*=.15) relative to control (see Figure 3, parts a and b). With regard to remission, moderately depressed participants had a greater response to Project: EVO (28/56, 50%) and iPST (39/79, 49%) than to the Health Tips arm (24/76, 32%; χ^2_2= 6.46, *P*=.04). We found no difference between treatment groups at week 4 (χ^2_2=0.84, *P*=.66) or week 8 (χ^2_2=1.79, *P*=.41).

Disability Outcomes

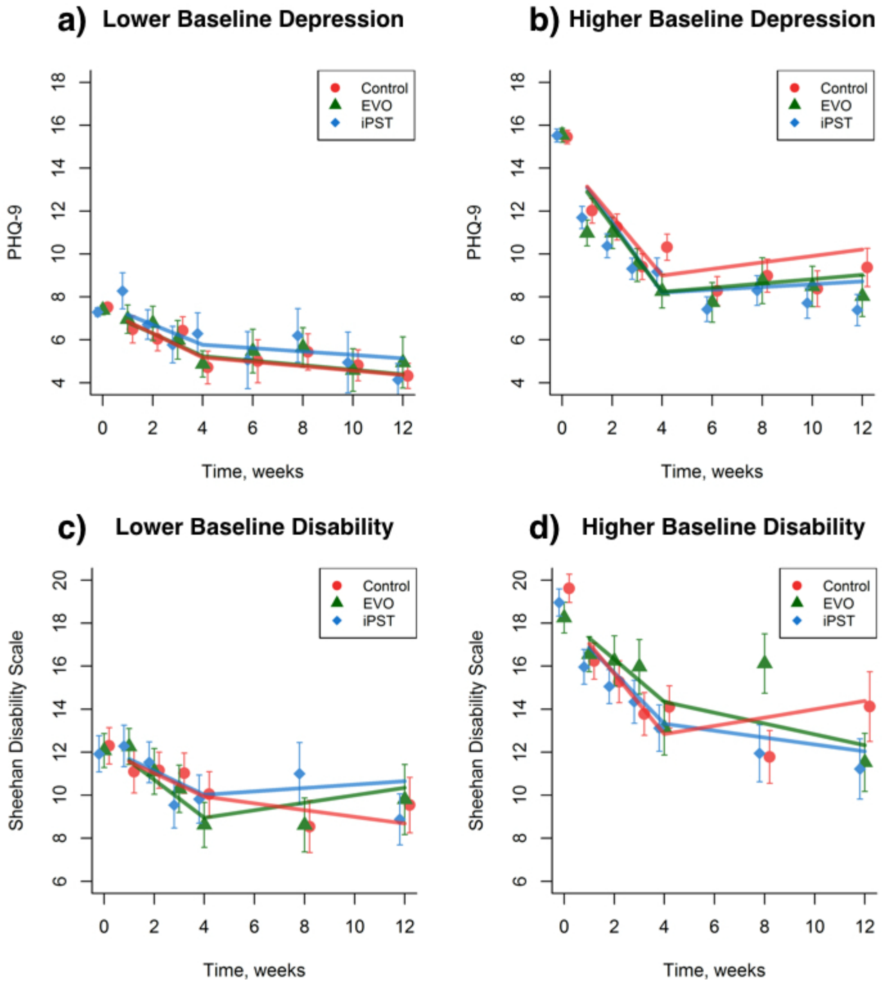
Disability Symptom Severity

Disability decreased an average of 0.67 points per week during weeks 1-4 and did not significantly change from weeks 4-12 (see Figure 2 d). The Project: EVO and iPST groups' disability did not significantly differ from that of controls at week 4 or week 8 or in the rates of change over time (see Table 3).

Outcomes by Baseline Disability Severity

A total of 159 participants were classified as having mild disability (SDS ≤15) and 237 with moderate disability at baseline (SDS >15). For the both subgroups, Project: EVO and iPST did not significantly differ from the control condition at week 4 or week 8. With regard to remission, moderately disabled participants showed no difference at weeks 4 and 8 for the Project: EVO (13/42, 31%; and 7/30, 23%), iPST (20/54, 37%; and 16/45, 36%), and Health Tips conditions (15/61, 25%; and 18/44, 41%; χ^2_2=2.09 and 2.48, respectively, *P*=.35 and .29, respectively; see Figure 3 c). For the mildly disabled group, Project: EVO yielded higher rates of recovery at 4 weeks (18/34, 53%) compared with the Health Tips group (9/38, 24%; χ^2_1=5.36, *P*=.02), with similar recovery observed between the iPST (14/40, 35%) and Health Tips arms (χ^2_1=0.72, *P*=.40; see Figure 3 d), with no group differences observed at week 8 (χ^2_2=1.31, *P*=.52).

Figure 3. Changes in 9-item Patient Health Questionnaire (PHQ-9) and Sheehan Disability Scale scores moderated by baseline depression and by symptom severity for each intervention. (a) Individuals with lower baseline depression by group. (b) Individuals with higher baseline depression by group. (c) Individuals with lower baseline anxiety by group. (d) Individuals with higher baseline anxiety by group. Control: Health Tips; EVO: Project: EVO; iPST: problem-solving therapy app. Error bars represent +/- 1 SE.



Discussion

Principal Findings

To our knowledge, this is the first fully remote randomized clinical field trial of mobile apps for depression in a national sample in the United States. Given the increased interest of health care organizations in the potential of mobile technology to address service gaps for mental health [45], the data from this study provide important information on the impact apps can have in depression care. Our findings suggest that apps designed to engage cognitive correlates of depression had the strongest effect on depressed mood for people with more moderate levels of depression.

As has been found in smaller-scale studies, people who download mood apps tend to use these tools as intended for no more than 2 weeks [21,46]. Despite poor adherence in app use, the benefits seem to be positive in terms of mood and functioning. While we cannot rule out that the outcomes were not simply regression to the mean for the full sample, it does appear that for those who are more significantly impaired, apps that are designed to target specific cognitive deficits implicit in depressive disorders, in this case cognitive control, are more effective than our control intervention. These findings, coupled with data from smaller proof-of-concept studies of the impact of apps on mood [17,47], suggest that for some people with mild to moderate depression, mobile apps could serve as an alternative means of treatment, particularly for those where mental health resources are scarce (eg, ethnic minorities). This was evidenced by the fact that our sample was much more representative of the US population than is typically found in mental health settings across the United States, with service utilization among ethnic/racial groups here being comparable with use rates in the United States [48]. However, these (and all) interpretations should be reviewed with caution given that, while this was an RCT, it was primarily conducted as a feasibility trial. Thus, this was not a mechanistic trial specifically designed to assess and compare the efficacy of these interventions, but rather to provide methodological insights for future work in this space.

Interpretation and Limitations

While our findings showed a positive impact on depression and disability outcomes, half of the enrolled participants never downloaded their assigned app despite having completed eligibility screens, consent forms, and baseline assessment. This is not an uncommon phenomenon in research of this nature: for example, nearly all self-guided Internet-based studies experience dropout rates as high as 90% very early in the study timeline [49], and a recent RCT comparing a mood app versus a control app had an 82% dropout rate [50]. Although our final sample was large enough to test the effects of apps on outcomes, the findings should be interpreted cautiously. For the mildly depressed subsample (those individuals with a PHQ-9 between 5 and 10), changes in mood could be attributed to regression to the mean, as spontaneous recovery with such mild depression is not uncommon. However, it is rare for people with a score of ≥15 on the PHQ-9 to simply "get better" (ie, regress to the mean), with our findings in the moderately depressed subsample demonstrating a significant difference between interventions. Furthermore, given the variation in outcomes based on symptom severity, and recent data finding that people with PHQ-9 scores of less than 10 do not have a clinical depression [22], we do not recommend that depression outcomes studies recruit participants with such mild presentations, unless the sample is very large (>10,000).

In addition to the high early study dropout rate, most participants did not use their assigned intervention apps as instructed, a ubiquitous effect across apps even given that these apps differed in content, user experience, and other features. This pattern of use mirrors other field trials of mental health apps that report app use typically wanes over the course of 2 weeks [51]. While limited adherence restricts our ability to test the efficacy of these interventions when used as designed, our data provide useful insight into how individuals typically interact with self-guided treatment apps. Indeed, the participants in this study were relatively tech savvy compared with those less technologically inclined, and we would expect to see some differences comparing these 2 groups, especially in the likelihood of using such apps for treatment purposes. For those less technologically inclined (or even those in this study), having a personal connection with a coach or other peers may be a critical element to encourage longer-term app use. Additionally, future research should also investigate the utility of apps that dynamically shift as user goals shift. The human-computer interaction field has recently demonstrated that user needs and interest in app-supported care vary over time, and engagement with app-based care may hinge on the ability of apps to dynamically adjust to the users' needs and interests [52].

Mobile apps still have a potential place in the treatment of mood disorders. Adherence to these tools, particularly when delivered as a self-guided treatment, is a problem, and methods for improving adherence should be explored.

Acknowledgments

Support for this research was provided by the National Institute of Mental Health (PAA R34-MH100466, T32MH0182607, K24MH074717). PAA had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. The authors thank M Brander (University of California, San Francisco, UCSF) and A Bodepudi (UCSF) for assistance in data monitoring; M Gross (UCSF) and J Camire (Ekho.me) for their help in participant recruitment; J Steinmetz (Ekho.me) for database architecture; D Ziegler (UCSF) for helping with app deployment; D Albert (UCSF) for assistance in Web design; C Catledge (UCSF) for administrative oversight; A Piper, E Martucci, S Kellogg, J Bower, M Omerick, and the entire Akili Interactive team, as well as I Elson, L Kaye, S Goobich, and the rest of the Ginger.io team, for helping with data collection and partnering with us on this project. We thank A Brandes-Aitken for assistance with supplemental cognitive analyses. The authors also would especially like to thank all the participants whose time and efforts made this work possible.

Conflicts of Interest

AG is cofounder, chief science advisor, and shareholder of Akili Interactive Labs, a company that develops cognitive training software. AG has a patent pending for a game-based cognitive training intervention, "Enhancing cognition in the presence of distraction and/or interruption," on which the cognitive training app (Project: EVO) that was used in this study was based. No other author has any conflict of interest to report.

Multimedia Appendix 1

Supplemental materials: Follow-up rates; eTable 1 (PHQ-9 by week as a function of Group Adherence for each active intervention); Adaptive Cognitive Evaluation (ACE); Cognitive Therapy (EVO); Problem Solving Therapy; Health Tips; Expectancy; Perceived Participant Burden.

[PDF File (Adobe PDF File), 409KB - jmir_v18i12e330_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist V1.6.1.

[PDF File (Adobe PDF File), 720KB - jmir_v18i12e330_app2.pdf]

References

1. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, National Comorbidity Survey Replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA 2003 Jun 18;289(23):3095-3105. [doi: 10.1001/jama.289.23.3095] [Medline: 12813115]
2. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
3. Bell AC, D'Zurilla TJ. Problem-solving therapy for depression: a meta-analysis. Clin Psychol Rev 2009 Jun;29(4):348-353. [doi: 10.1016/j.cpr.2009.02.003] [Medline: 19299058]
4. Cuijpers P, Geraedts AS, van Oppen P, Andersson G, Markowitz JC, van Straten A. Interpersonal psychotherapy for depression: a meta-analysis. Am J Psychiatry 2011 Jun;168(6):581-592 [FREE Full text] [doi: 10.1176/appi.ajp.2010.10101411] [Medline: 21362740]
5. Greenberg DB. Barriers to the treatment of depression in cancer patients. J Natl Cancer Inst Monogr 2004(32):127-135. [doi: 10.1093/jncimonographs/lgh019] [Medline: 15263054]
6. Hirschfeld RM, Keller MB, Panico S, Arons BS, Barlow D, Davidoff F, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. JAMA 1997;277(4):333-340. [Medline: 9002497]
7. Fortney JC, Harman JS, Xu S, Dong F. The association between rural residence and the use, type, and quality of depression care. J Rural Health 2010;26(3):205-213. [doi: 10.1111/j.1748-0361.2010.00290.x] [Medline: 20633088]
8. Muller I, Yardley L. Telephone-delivered cognitive behavioural therapy: a systematic review and meta-analysis. J Telemed Telecare 2011;17(4):177-184. [doi: 10.1258/jtt.2010.100709] [Medline: 21357672]
9. Mohr DC, Vella L, Hart S, Heckman T, Simon G. The effect of telephone-administered psychotherapy on symptoms of depression and attrition: a meta-analysis. Clin Psychol (New York) 2008;15(3):243-253 [FREE Full text] [doi: 10.1111/j.1468-2850.2008.00134.x] [Medline: 21369344]
10. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
11. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Aug;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
12. Anderson M. Technology device ownership. Washington, DC: Pew Research Center; 2015. URL: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/ [accessed 2016-12-12] [WebCite Cache ID 6mhsOhgRL]
13. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
14. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
15. Shen N, Levitan M, Johnson A, Bender JL, Hamilton-Page M, Jadad AA, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
16. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
17. 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]
18. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
19. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
20. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
21. Rosa C, Campbell AN, Miele GM, Brunner M, Winstanley EL. Using e-technologies in clinical trials. Contemp Clin Trials 2015 Nov;45(Pt A):41-54. [doi: 10.1016/j.cct.2015.07.007] [Medline: 26176884]
22. 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]

23. Sheehan KH, Sheehan DV. Assessing treatment effects in clinical trials with the discan metric of the Sheehan Disability Scale. Int Clin Psychopharmacol 2008 Mar;23(2):70-83. [doi: 10.1097/YIC.0b013e3282f2b4d6] [Medline: 18301121]
24. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
25. Unützer J, Katon W, Callahan CM, Williams JW, Hunkeler E, Harpole L, IMPACT Investigators. Improving Mood-Promoting Access to Collaborative Treatment. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA 2002 Dec 11;288(22):2836-2845. [Medline: 12472325]
26. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [Medline: 9738608]
27. Paulus M. Cognitive control in depression and anxiety: out of control? Curr Opin Behav Sci 2015:003-020 [FREE Full text]
28. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, et al. Video game training enhances cognitive control in older adults. Nature 2013 Sep 5;501(7465):97-101 [FREE Full text] [doi: 10.1038/nature12486]
29. Anguera J, Gunning F, Arean P. Improving late life depression and cognitive control through the use of therapeutic video game technology: A proof-of-concept randomized trial. Depression Anxiety 2016:1-11 (forthcoming).
30. Malouff JM, Thorsteinsson EB, Schutte NS. The efficacy of problem solving therapy in reducing mental and physical health problems: a meta-analysis. Clin Psychol Rev 2007 Jan;27(1):46-57. [doi: 10.1016/j.cpr.2005.12.005]
31. Cuijpers P, van Straten A, Warmerdam L. Problem solving therapies for depression: a meta-analysis. Eur Psychiatry 2007 Jan;22(1):9-15. [doi: 10.1016/j.eurpsy.2006.11.001]
32. Karkar R, Zia J, Vilardaga R, Mishra SR, Fogarty J, Munson SA, et al. A framework for self-experimentation in personalized health. J Am Med Inform Assoc 2016 May;23(3):440-448. [doi: 10.1093/jamia/ocv150] [Medline: 26644399]
33. Hyewon S, Shahriaree N, Hekler E, Kientz J. Developing validating the User Burden Scale: a tool for assessing user burden in computing systems. 2016 Presented at: CHI Conference on Human Factors in Computing Systems. Santa Clara, CA, USA; May 7-12, 2016; Santa Clara, CA, USA p. 7-12. [doi: 10.1145/2858036.2858448]
34. Sheehan DV, Harnett-Sheehan K, Raj B. The measurement of disability. Int Clin Psychopharmacol 1996 Jun;11 Suppl 3:89-95. [Medline: 8923116]
35. Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C. A tutorial on count regression and zero-altered count models for longitudinal substance use data. Psychol Addict Behav 2013 Mar;27(1):166-177 [FREE Full text] [doi: 10.1037/a0029508]
36. Hallgren KA, Witkiewitz K. Missing data in alcohol clinical trials: a comparison of methods. Alcohol Clin Exp Res 2013 Dec;37(12):2152-2160 [FREE Full text] [doi: 10.1111/acer.12205]
37. Curran PJ, Obeidat K, Losardo D. Twelve frequently asked questions about growth curve modeling. J Cogn Dev 2010;11(2):121-136. [doi: 10.1080/15248371003699969] [Medline: 21743795]
38. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs. technician assistance. PLoS One 2010;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939]
39. Richards DA, Suckling R. Improving access to psychological therapies: phase IV prospective cohort study. Br J Clin Psychol 2009 Nov;48(Pt 4):377-396. [doi: 10.1348/014466509X405178]
40. Hedeker D, Gibbons R. Application of random-effects pattern-mixture models for missing data in longitudinal studies. Psychol Methods 1997;2(1):64-78.
41. Wright CC, Sim J. Intention-to-treat approach to data from randomized controlled trials: a sensitivity analysis. J Clin Epidemiol 2003 Sep;56(9):833-842. [Medline: 14505767]
42. Armijo-Olivo S, Warren S, Magee D. Intention to treat analysis, compliance, drop-outs and how to deal with missing data in clinical research: a review. Phys Ther Rev 2009;14(1):36-49. [doi: 10.1179/174328809X405928]
43. Yuan Y. Sensitivity analysis in multiple imputation for missing data. Cary, NC: SAS Institute Inc; 2014. URL: https://support.sas.com/resources/papers/proceedings14/SAS270-2014.pdf [accessed 2016-12-12] [WebCite Cache ID 6mhu3cE7y]
44. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Software 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
45. Banos O, Villalonga C, Garcia R, Saez A, Damas M, Holgado-Terriza JA, et al. Design, implementation and validation of a novel open framework for agile development of mobile health applications. Biomed Eng Online 2015;14 Suppl 2:S6 [FREE Full text] [doi: 10.1186/1475-925X-14-S2-S6] [Medline: 26329639]
46. Lee J, Nguyen AL, Berg J, Amin A, Bachman M, Guo Y, et al. Attitudes and preferences on the use of mobile health technology and health games for self-management: interviews with older adults on anticoagulation therapy. JMIR Mhealth Uhealth 2014;2(3):e32 [FREE Full text] [doi: 10.2196/mhealth.3196] [Medline: 25098413]
47. Burns MN, Montague E, Mohr DC. Initial design of culturally informed behavioral intervention technologies: developing an mHealth intervention for young sexual minority men with generalized anxiety disorder and major depression. J Med Internet Res 2013;15(12):e271 [FREE Full text] [doi: 10.2196/jmir.2826] [Medline: 24311444]
48. Results from the 2015 National Survey on Drug Use and Health: detailed tables. Table 8.38B – Received outpatient mental health treatment/counseling in past year among persons aged 18 or older, by past year level of mental illness, demographic

characteristics, and socioeconomic characteristics: percentages, 2014 and 2015. Rockville, MD: Substance Abuse and Mental Health Services Administration. URL: http://www.samhsa.gov/data/sites/default/files/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015/NSDUH-DetTabs-2015.htm#tab8-38b [accessed 2016-12-15] [WebCite Cache ID 6mmGOc39h]

49. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of Web-based interventions: a design research viewpoint. J Med Internet Res 2015;17(7):e172 [FREE Full text] [doi: 10.2196/jmir.4201] [Medline: 26163456]
50. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of superbetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
51. Kuhn E, Greene C, Hoffman J, Nguyen T, Wald L, Schmidt J, et al. Preliminary evaluation of PTSD Coach, a smartphone app for post-traumatic stress symptoms. Mil Med 2014 Jan;179(1):12-18. [doi: 10.7205/MILMED-D-13-00271] [Medline: 24402979]
52. Li I, Dey A, Forlizzi J. A stage-based model of personal informatics systems. 2010 Presented at: SIGCHI Conference on Human Factors in Computing Systems (CHI 10); April 10-15, 2010; Atlanta, GA, USA p. 10-15. [doi: 10.1145/1753326.1753409]

Abbreviations

AUDIT-C: Alcohol Use Disorders Identification Test
CONSORT: Consolidated Standards of Reporting Trials
GAD-7: Generalized Anxiety Disorder 7-item scale
HIPAA: Health Insurance Portability and Accountability Act
IMPACT: Improving Mood-Promoting Access to Collaborative Treatment
OR: odds ratio
PHQ-9: 9-item Patient Health Questionnaire
RCT: randomized controlled trial
RR: rate ratio
SDS: Sheehan Disability Scale
SMS: short message service

Edited by G Eysenbach; submitted 11.08.16; peer-reviewed by K Kauppi, S Schueller, D Ben-Zeev; comments to author 14.09.16; revised version received 14.11.16; accepted 30.11.16; published 20.12.16.

Please cite as:
Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, Anguera JA
The Use and Effectiveness of Mobile Apps for Depression: Results From a Fully Remote Clinical Trial
J Med Internet Res 2016;18(12):e330
URL: http://www.jmir.org/2016/12/e330/
doi:10.2196/jmir.6482
PMID:27998876

Original Paper

Social Media Use Among Living Kidney Donors and Recipients: Survey on Current Practice and Potential

Abby Swanson Kazley[1*], PhD; Bashir Hamidi[2*], MPH; Wendy Balliet[3*], PhD; Prabhakar Baliga[2*], MD

[1]Department of Health Care Leadership and Management, Medical University of South Carolina, Charleston, SC, United States

[2]Department of Surgery, Medical University of South Carolina, Charleston, SC, United States

[3]Department of Psychiatry and Behavioral Medicine, Medical University of South Carolina, Charleston, SC, United States

[*]all authors contributed equally

Corresponding Author:
Abby Swanson Kazley, PhD
Department of Health Care Leadership and Management
Medical University of South Carolina
151 Rutledge Ave
Charleston, SC, 29425
United States
Phone: 1 843 792 0012
Fax: 1 843 792 3327
Email: swansoaj@musc.edu

Abstract

Background: In the United States, there is a national shortage of organs donated for transplant. Among the solid organs, most often kidneys are donated by living donors, but the lack of information and complicated processes limit the number of individuals who serve as living kidney donors. Social media can be a tool for advocacy, educating the public about the need, process, and outcomes of live kidney donors, yet little is known about social media use by kidney transplant patients.

Objective: The purpose of this study was to examine the social media use of potential kidney transplant patients and their willingness to use social media and their networks to advocate and educate about living kidney donation.

Methods: Using a validated survey, we modified the instrument to apply to the patient population of interest attending the Medical University of South Carolina, Charleston, SC, USA. The questions on the survey inquired about current social media use, sites visited, frequency and duration of social media use, and willingness to use social media to share the need for living kidney donors. We asked patients who had received a transplant and those awaiting a transplant to complete the survey during an office visit. Participation was voluntary.

Results: A total of 199 patients completed the survey. Approximately half of all kidney transplant patients surveyed used social media (104/199, 52.3%), and approximately one-third (66/199, 33.2%) had more than 100 friends in their social media network. Facebook was the most popular site, and 51% (102/199) reported that they would be willing to post information about living kidney donation on their social networks. More than a quarter of the sample (75/199, 37.7%) had posted about their health status in the past.

Conclusions: Social media holds great promise for health-related education and awareness. Our study shows the current social media use of kidney transplant patients. In turn, such information can be used to design interventions to ensure appropriate decision making about live kidney donation. Transplant programs can help increase the number of living donors by providing guidance to kidney transplant patients in how to use social media, to be advocates, and to provide information about living kidney donation to their social network.

(J Med Internet Res 2016;18(12):e328) doi:10.2196/jmir.6176

KEYWORDS

living kidney donation; kidney transplant; social media

Introduction

For patients with end-stage renal disease, kidney transplant is the best available treatment, improving longevity of life (average 10 years), optimizing quality of life, and costing significantly less than dialysis treatments. One of the biggest challenges in the US transplantation community (though this problem exists worldwide) is the growing gap between the increasing demand and the number of procured organs [1,2]. In the United States, there are more than 100,000 candidates on the kidney transplant waitlist. However, only 17,878 kidney transplants were performed in the United States in 2015 [3].

Recent strategies to reduce waitlist time-to-transplant have focused on increasing living donor kidney transplant (LDKT) [4-6]. LDKT is associated with improved graft function, better patient survival, and lower health-related financial costs [7]. Although the risks of donation are minimal (eg, potential hypertension, hernia, and <0.01% mortality), there has been an overall decrease in rates of LDKT [4,8]. Barriers to living kidney donation include difficulty communicating about the need for kidney donation with potential donors, worry about financial strain for the donor, fear of health risks to the donor, and a general lack of knowledge about the kidney donation and transplant process. Patients and transplantation professionals must continue to seek new ways to promote awareness and provide education about the opportunity to be a living donor.

Social media has proven to be an effective tool for increasing awareness about various chronic illnesses and has been used to promote positive health behavior change for certain health-related topics. The literature supports the utility of social media campaigns to promote awareness about cancer and to reduce rates of smoking, alcohol abuse, and use of illicit drugs [9-11]. Further, social media has been shown to reinforce positive behaviors such as healthy eating, regular exercise, and frequent preventive medical screenings [12]. One of the notable campaigns using social media, with a goal of increasing the number of donor registrations, has been a remarkable collaborative effort between the transplantation team at the Johns Hopkins University School of Medicine and Facebook that resulted in the addition of a feature on Facebook that both raised awareness about donation and recruited participants into donor registries. Known as the “Facebook effect,” this effort was vastly successful, increasing the number of new donor registrations by approximately 21-fold by the day after the implementation and by a 5.8-fold increase over baseline over a 13-day period [13].

These innovative health-promoting social media initiatives highlight the great potential to increase awareness and provide education about living kidney donation by capitalizing on the influence of social media in the United States. Social media can be a tool used in creative ways to educate hard-to-reach populations. With the worldwide use of social media and its structure of social group communication, there can be opportunities for patients to use their social media networks as a way to increase LDKT, with the goals of reducing time on the waitlist and improving quality of life both before and after transplant. This can be accomplished by raising awareness and educating others about the need for living kidney donation. However, the literature is lacking in understanding whether or how potential transplant patients use social media. The purpose of this study was to examine and describe social media use by individuals in South Carolina who have been referred for a kidney transplant. This study allowed us to learn about patients’ current practices with social media and evaluate the potential for future interventions that could raise awareness of the opportunity for living kidney donation through the use of social media.

Methods

We used a cross-sectional design to examine patients’ current use of social media for health-related and nonhealth-related topics. Additionally, we assessed patients’ perceptions of social media use for specific kidney transplant health-related activities. Study data were collected and managed using REDCap electronic data capture tools hosted at the Medical University of South Carolina, Charleston, SC, USA [14]. A research associate administered approximately 200 survey interviews to both kidney transplant candidates and kidney transplant recipients who were seen at the Medical University of South Carolina Renal Access Clinic. Kidney transplant candidates were defined as patients who did not have a kidney transplant prior to administration of the survey and who were attending clinic as a pretransplant requirement to be placed on our kidney transplant waitlist. Kidney transplant recipients were defined as those who underwent a kidney transplant and were attending clinic visits as a routine follow-up or due to an acute issue.

Data collection was not identified to any particular patients and data were only examined in the aggregate. Patients provided verbal consent during their clinic visit and, if they were willing to participate, were interviewed using the survey questions. Data were collected between May and September 2015, and the study was approved by the institutional review board of the Medical University of South Carolina.

Using a validated survey with permission from PricewaterhouseCoopers LLP (New York, NY, USA), we sought to identify social media use in our sample [15]. Due to the specific aims of our study, we added questions to the survey to gauge the willingness of patients to share transplant-related information. The added questions were written with expert input and pilot tested with a sample of 15. Descriptive analyses, including frequencies, means, and ranges, were calculated for all the variables. We computed 1-way analysis of variance and chi-square analyses to determine any potential differences among transplant status, and social media use and demographic variables (eg, age, income, education, sex), using IBM SPSS version 23 (IBM Corporation).

Results

The total number of patients surveyed was 199. Of these 199 patients, a majority (n=115, 57.8%) were male. A majority of the patients were kidney transplant recipients (n=133), and 66 were kidney transplant candidates. In terms of educational level attained, 29 patients had not graduated from high school, and

8 had an eighth grade education or less. More than one-third had graduated from high school, and more than half had at least some college education. The majority of the sample ranged in age from 45 to 64 years, while 21.1% (n=42) were older than 65 years, and 24.1% (n=48) were less than 45 years of age. Table 1 lists these results.

More than half of the sample (n=104, 52.3%) reported that they used social media sites on the Internet. Due to small sample sizes within groups, we transformed age to include only 2 groups (18-44 and ≥44 years), and we collapsed education to comprise 3 groups (less than a high school diploma, high school degree and some college, 2-year college degree or higher). Pearson chi-square analyses comparing potential differences between transplant status (pre- or post-) and social media use revealed no significant difference between those who had a transplant and those who had not in their tendency to use social media (χ^2_{12}=.023, P>.99) . No significant differences were found between transplant status and income, sex, age, or education.

In terms of social media use, there was a significant difference between age and social media use, with younger people more likely than older adults to use social media (χ^2_8=26.15, P<.001). We found no significant differences between social media use and any other demographic variables.

Nearly half (n=96) of the patients sampled had been using social media for less than a month, whereas 31.2% (n=62) had been on social media for over 4 years. Of those who used social media, 24.1% (n=48) reported using social media for 5 hours or less per week. Given the purpose of the study, the size of the social media networks was of particular interest. Of those on social media, 33.2% (n=66) had over 100 “friends” in their social network. Table 2 presents these results.

The most frequently used social media site was Facebook, followed by Instagram, Twitter, and Google+. None of the participants used Reddit, Foursquare, or LiveJournal. Table 3 presents these results.

Table 1. Descriptive data of kidney transplant candidates and kidney transplant recipients, Medical University of South Carolina Renal Access Clinic, May-September 2015 (N=199).

Characteristics	n (%)
Sex	
Male	115 (57.8)
Female	84 (42.2)
Transplant status	
Pretransplant	66 (33.2)
Posttransplant	133 (66.8)
Age range (years)	
18-24	5 (2.5)
25-34	14 (7.0)
35-44	29 (14.6)
45-54	54 (27.1)
55-64	55 (27.6)
≥65 years	42 (21.1)
Educational level attained	
≤8th grade	8 (4.0)
Some high school	21 (10.6)
Graduated from high school	69 (34.7)
Some college	41 (20.6)
2-year college	25 (12.6)
≥4-year college	35 (17.6)

Table 2. Summary of social media use by kidney transplant candidates and kidney transplant recipients.

Survey questions	n (%)
Do you use social networking sites on the Internet?	
Yes	104 (52.3)
No	95 (47.7)
How long have you been on social media?	
<1 month	96 (48.2)
1-6 months	3 (1.5)
6-12 months	9 (4.5)
1-2 years	10 (5.0)
2-4 years	18 (9.0)
≥5 years	62 (31.2)
Missing	1 (0.5)
Total	199
How many hours per week do you spend on social media?	
N/A[a]	97 (48.7)
0-5 hours	48 (24.1)
6-10 hours	29 (14.6)
11-20 hours	10 (5.0)
21-30 hours	5 (2.5)
≥31 hours	9 (4.5)
Missing	1 (0.5)
Total	199
How many “friends” do you have on social media?	
0-10	6 (3.0)
11-50	17 (8.5)
51-100	15 (7.5)
101-250	22 (11.1)
≥251	44 (22.1)
Total	104
N/A	95

[a] N/A: not applicable.

Table 3. Social media of choice among kidney transplant candidates and kidney transplant recipient.

Social media site used	n
Facebook	102
Google+	16
Tumblr	2
Twitter	18
Instagram	20
Pinterest	13
Reddit	0
Vine	3
LiveJournal	0
Foursquare	0
LinkedIn	12
Myspace	2
Snapchat	8
Other	3

Once we established that the majority of our sample used social media, we sought to gauge their comfort and willingness to use it to promote health-related causes. Results indicated that 25.1% (n=50) of patients questioned had used social media to post health-related activities, while 37.7% (n=75) had posted about nonhealth-related activities. Importantly, 35.7% of participants (n=71) reported that they would be willing to share information about their health through social media. And nearly half (n=94, 47.2%) of patients indicated that they would be willing to post information about their kidney disease on social media. Survey results revealed that several patients or their families had already used social media to promote their need for living kidney donation: 14 patients (7%) had a family or friend post about their need for a living kidney donor, 3 patients had asked for a living kidney donation through social media, and 6 (3.0%) had both posted themselves and had a friend or relative post for them. In total, 23 individuals (11.6%) had used social medial to ask for a living kidney donation. Table 4 presents these results.

Table 4. Health-related social media use and willingness to use it to promote health-related causes.

Survey questions	n (%)
Have you posted health-related activities using social media?	
Yes	50 (25.1)
No	54 (27.1)
Total	104
N/A[a]	95 (47.7)
Have you posted nonhealth-related activities on social media?	
Yes	75 (37.7)
No	28 (14.1)
Total	103
N/A	96 (48.2)
Would you be willing to share information about your health through social media?	
Yes	71 (35.7)
No	33 (16.6)
Total	104
N/A	95 (47.7)
Would you be willing to post information about your kidney disease on social media?	
Yes	94 (47.2)
No	41 (20.6)
Total	135 (67.8)[b]
Missing	64 (32.2)
Did you or a family or friend share information on social media about your need for a kidney transplant?	
Neither asked	74 (37.2)
Patient asked	3 (1.5)
Family or friend asked	14 (7.0)
Both asked	6 (3.0)
Total	97 (48.7)
Missing	102 (51.3)

[a] N/A: not applicable.

[b] Includes people who said “no” to a history of sharing.

Discussion

The utility of social media holds great potential for connecting individuals with others, as well as for providing health-related education. Results from this study revealed that social media is used by many patients in this population who are in need of, or already had, a kidney transplant. Importantly, responses indicate that almost half of patients surveyed would be willing to use social media to share information about their kidney disease. Younger adults are more likely to use social media, which suggests that older adults (≥45 years) could be targeted as a population of intervention to educate about social media use. The results demonstrate the vast potential that social media has for raising awareness about chronic illness and providing health care-related education.

From a transplant-specific perspective, we have shown that some patients and family members have already begun using social media to promote the need for living kidney donors. Moreover, many patients who haven’t yet used social media in this way indicated a willingness to do so. Because social media has been effective in raising awareness for other health-related topics, the global potential in using it as a tool to raise awareness about renal disease and the prospect of living donor transplants is immeasurable. One of the significant barriers to living kidney donation is the kidney transplant candidate’s hesitation in initiating conversation with potential donors, family, and friends. Therefore, social media use offers an innovative and unique opportunity for transplant candidates to promote awareness about renal disease and the negative effect of various treatments (eg, dialysis), and about their need for a kidney, in a thoughtful and less anxiety-provoking way, thereby potentially increasing their network size of potential living kidney donors. Despite the sensitivity and privacy that sometimes accompany illness, results from this study suggest that a good number of patients would be willing to share information about their kidney disease using social media, yet only a small number have done so.

Interestingly, kidney disease patients sampled in this study used social media less than the general population of adults in the United States (65%) [16]. Results then also suggest that the kidney transplant population is actually underrepresented on social media. This indicates that patients may need guidance in using social media in general, and also to promote their need for a transplant. It is also possible that many of the patients who do use social media have not yet considered using it for this purpose.

Given the potential networking capability of social media, a well-presented case of an individual patient in need of a kidney transplant may raise awareness and lead individuals who know the patient, as well as altruistic donors, to come forward. Since Gallup data have shown that a majority of Americans support organ donation, there is reason to believe that individuals would be receptive to such targeted messages [17]. There have been a handful of already publicized cases of living donors who came forward after learning of the need of an individual patient through social media [18,19]. Furthermore, with the option of paired kidney exchanges and donation chains, the potential for donations to happen is even greater than if a single person comes forward to make a single donation. Given the abundant and growing need for transplantable organs, social media may provide a tool to increase awareness and promote growth in living kidney donation.

Future research may evaluate the use of a social media intervention to increase public awareness of the potential to serve as a living kidney donor and, subsequently, the number of living donors who are evaluated and have learned of the option of living kidney donation through a social media app. Such interventions will have to heed which social media sites are the most popular, as the goal would be to reach the widest audience possible. Transplant centers may consider providing coaching or examples of how social media messages should be written to be compelling, accurate, and informative. Although some patients have larger social media networks (ie, number of friends), others do not. Those who do not can still use the tool by asking their “connections” to also share their message, with the option to boost posts geographically.

Our study does have limitations. First, the data are from only a single transplant center and, thus, the results may not be representative of other geographical areas. Second, the results are self-reported and thus may include some responder bias. It was surprising that 48% of participants had only used social media for less than a month; this may suggest survey reading confusion. We also did not gather information about etiology of kidney disease or ethnicity, which could have added to our understanding of social media use in this population. Third, this was an exploratory study examining the potential of social media in living kidney transplant, and our results are not tied to any specific outcomes.

Acknowledgments

REDCap is supported by the South Carolina Clinical and Translational Research Institute, with an academic home at the Medical University of South Carolina, through National Institutes of Health National Center for Advancing Translational Sciences grant number UL1 TR000062.

Conflicts of Interest

None declared.

References

1. Anonymous. Preface. OPTN/SRTR 2013 annual data report. Am J Transplant 2015 Jan;15 Suppl 2:4-7 [FREE Full text] [doi: 10.1111/ajt.13221] [Medline: 25626349]
2. Langone AJ, Helderman JH. Disparity between solid-organ supply and demand. N Engl J Med 2003 Aug 14;349(7):704-706. [doi: 10.1056/NEJMe038117] [Medline: 12917309]
3. Organ Procurement and Transplantation Network. National Data. Richmond, VA: US Department of Health and Human Services; 2016. URL: https://optn.transplant.hrsa.gov/data/view-data-reports/national-data/# [accessed 2016-12-09] [WebCite Cache ID 6mdILt2na]
4. Axelrod DA, McCullough KP, Brewer ED, Becker BN, Segev DL, Rao PS. Kidney and pancreas transplantation in the United States, 1999-2008: the changing face of living donation. Am J Transplant 2010 Apr;10(4 Pt 2):987-1002 [FREE Full text] [doi: 10.1111/j.1600-6143.2010.03022.x] [Medline: 20420648]
5. Rodrigue JR, Pavlakis M, Danovitch GM, Johnson SR, Karp SJ, Khwaja K, et al. Evaluating living kidney donors: relationship types, psychosocial criteria, and consent processes at US transplant programs. Am J Transplant 2007 Oct;7(10):2326-2332 [FREE Full text] [doi: 10.1111/j.1600-6143.2007.01921.x] [Medline: 17845566]
6. Abecassis M, Bartlett ST, Collins AJ, Davis CL, Delmonico FL, Friedewald JJ, et al. Kidney transplantation as primary therapy for end-stage renal disease: a National Kidney Foundation/Kidney Disease Outcomes Quality Initiative (NKF/KDOQITM) conference. Clin J Am Soc Nephrol 2008 Mar;3(2):471-480 [FREE Full text] [doi: 10.2215/CJN.05021107] [Medline: 18256371]

7. Smith C, Woodward RS, Cohen DS, Singer GG, Brennan DC, Lowell JA, et al. Cadaveric versus living donor kidney transplantation: a Medicare payment analysis. Transplantation 2000 Jan 27;69(2):311-314. [Medline: 10670645]
8. Rodrigue JR, Kazley AS, Mandelbrot DA, Hays R, LaPointe RD, Baliga P, American Society of Transplantation. Living donor kidney transplantation: overcoming disparities in live kidney donation in the US: recommendations from a consensus conference. Clin J Am Soc Nephrol 2015 Sep 04;10(9):1687-1695 [FREE Full text] [doi: 10.2215/CJN.00700115] [Medline: 25883072]
9. Bala M, Strzeszynski L, Cahill K. Mass media interventions for smoking cessation in adults. Cochrane Database Syst Rev 2008(1):CD004704. [doi: 10.1002/14651858.CD004704.pub2] [Medline: 18254058]
10. Farrelly MC, Nonnemaker J, Davis KC, Hussin A. The Influence of the National truth campaign on smoking initiation. Am J Prev Med 2009 May;36(5):379-384. [doi: 10.1016/j.amepre.2009.01.019] [Medline: 19211213]
11. Staccini P, Fernandez-Luque L. Health social media and patient-centered care: buzz or evidence? Findings from the Section “Education and Consumer Health Informatics” of the 2015 Edition of the IMIA Yearbook. Yearb Med Inform 2015 Aug 13;10(1):160-163 [FREE Full text] [doi: 10.15265/IY-2015-032] [Medline: 26293862]
12. 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]
13. Cameron AM, Massie AB, Alexander CE, Stewart B, Montgomery RA, Benavides NR, et al. Social media and organ donor registration: the Facebook effect. Am J Transplant 2013 Aug;13(8):2059-2065 [FREE Full text] [doi: 10.1111/ajt.12312] [Medline: 23777475]
14. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
15. Health Research Institute. Social media “likes” healthcare: from marketing to social business. New York, NY: PricewaterhouseCoopers LLP; 2015. URL: http://www.pwc.com/us/en/health-industries/health-research-institute/publications/health-care-social-media.html [accessed 2016-12-07] [WebCite Cache ID 6maJRMz1G]
16. Pew Research Center. Internet use over time. Washington, DC: Pew Research Center; 2016. URL: http://www.pewinternet.org/data-trend/internet-use/internet-use-over-time/ [accessed 2016-12-14] [WebCite Cache ID 6ml33RFNu]
17. Siegel JT, Navarro MA, Tan CN, Hyde MK. Attitude-behavior consistency, the principle of compatibility, and organ donation: a classic innovation. Health Psychol 2014 Sep;33(9):1084-1091. [doi: 10.1037/hea0000062] [Medline: 24707844]
18. Gann C. Before donation, man writes ‘goodbye’ letters to his kidney: ‘good job these last 25 years’. ABC News. New York, NY: ABC News Internet Ventures; 2012 May 03. URL: http://abcnews.go.com/Health/kidney-donation-good-samaritan-donor-wrote-goodbye-letters/story?id=16264197 [accessed 2016-12-14] [WebCite Cache ID 6ml1VZAry]
19. Hartman S. Searching for a life-saving kidney: and finding the kindness of strangers. CBS News. New York, NY: CBS Interactive Inc; 2012 Sep 07. URL: http://www.cbsnews.com/news/searching-for-a-life-saving-kidney-and-finding-the-kindness-of-strangers/ [accessed 2016-12-14] [WebCite Cache ID 6ml14x1Gn]

Abbreviations

LDKT: living donor kidney transplant

Edited by G Eysenbach; submitted 08.06.16; peer-reviewed by M Leveridge, D Walker; comments to author 21.07.16; revised version received 19.09.16; accepted 07.10.16; published 20.12.16.

Please cite as:
Kazley AS, Hamidi B, Balliet W, Baliga P
Social Media Use Among Living Kidney Donors and Recipients: Survey on Current Practice and Potential
J Med Internet Res 2016;18(12):e328
URL: http://www.jmir.org/2016/12/e328/
doi:10.2196/jmir.6176
PMID:27998880

Original Paper

Applying Multiple Data Collection Tools to Quantify Human Papillomavirus Vaccine Communication on Twitter

Philip M Massey[1*], MPH, PhD; Amy Leader[2*], MPH, DrPH; Elad Yom-Tov[3*], MSc, PhD; Alexandra Budenz[1*], MA; Kara Fisher[4*], MPH; Ann C Klassen[1*], PhD

[1]Department of Community Health and Prevention, Dornsife School of Public Health, Drexel University, Philadelphia, PA, United States

[2]Division of Population Science, Department of Medical Oncology, Thomas Jefferson University, Philadelphia, PA, United States

[3]Microsoft Research Israel, Herzeliya, Israel

[4]Department of Biostatistics and Epidemiology, Dornsife School of Public Health, Drexel University, Philadelphia, PA, United States

[*]all authors contributed equally

Corresponding Author:
Philip M Massey, MPH, PhD
Department of Community Health and Prevention
Dornsife School of Public Health
Drexel University
Nesbitt Hall
3215 Market St
Philadelphia, PA, 19104
United States
Phone: 1 267 359 6067
Fax: 1 267 359 6109
Email: pmm85@drexel.edu

Abstract

Background: Human papillomavirus (HPV) is the most common sexually transmitted infection in the United States. There are several vaccines that protect against strains of HPV most associated with cervical and other cancers. Thus, HPV vaccination has become an important component of adolescent preventive health care. As media evolves, more information about HPV vaccination is shifting to social media platforms such as Twitter. Health information consumed on social media may be especially influential for segments of society such as younger populations, as well as ethnic and racial minorities.

Objective: The objectives of our study were to quantify HPV vaccine communication on Twitter, and to develop a novel methodology to improve the collection and analysis of Twitter data.

Methods: We collected Twitter data using 10 keywords related to HPV vaccination from August 1, 2014 to July 31, 2015. Prospective data collection used the Twitter Search API and retrospective data collection used Twitter Firehose. Using a codebook to characterize tweet sentiment and content, we coded a subsample of tweets by hand to develop classification models to code the entire sample using machine learning procedures. We also documented the words in the 140-character tweet text most associated with each keyword. We used chi-square tests, analysis of variance, and nonparametric equality of medians to test for significant differences in tweet characteristic by sentiment.

Results: A total of 193,379 English-language tweets were collected, classified, and analyzed. Associated words varied with each keyword, with more positive and preventive words associated with “HPV vaccine” and more negative words associated with name-brand vaccines. Positive sentiment was the largest type of sentiment in the sample, with 75,393 positive tweets (38.99% of the sample), followed by negative sentiment with 48,940 tweets (25.31% of the sample). Positive and neutral tweets constituted the largest percentage of tweets mentioning prevention or protection (20,425/75,393, 27.09% and 6477/25,110, 25.79%, respectively), compared with only 11.5% of negative tweets (5647/48,940; $P<.001$). Nearly one-half (22,726/48,940, 46.44%) of negative tweets mentioned side effects, compared with only 17.14% (12,921/75,393) of positive tweets and 15.08% of neutral tweets (3787/25,110; $P<.001$).

Conclusions: Examining social media to detect health trends, as well as to communicate important health information, is a growing area of research in public health. Understanding the content and implications of conversations that form around HPV vaccination on social media can aid health organizations and health-focused Twitter users in creating a meaningful exchange of ideas and in having a significant impact on vaccine uptake. This area of research is inherently interdisciplinary, and this study supports this movement by applying public health, health communication, and data science approaches to extend methodologies across fields.

(J Med Internet Res 2016;18(12):e318) doi:10.2196/jmir.6670

KEYWORDS

HPV vaccine; Twitter; communication methods; content analysis; data mining

Introduction

Human papillomavirus (HPV) is the most common sexually transmitted infection in the United States [1]. In the United States, approximately 79 million people are infected with HPV, and 14 million will become newly infected each year [2,3]. Although many infections will resolve without serious consequences, HPV infection has been causally linked to cervical and anal cancers, as well as genital warts. Several HPV vaccines are licensed in the United States that protect against strains of HPV most associated with cervical cancer in females and genital warts in males [4]. Thus, HPV vaccination has become an important component of adolescent preventive health care. According to the US Centers for Disease Control and Prevention, in the United States, HPV vaccination rates have steadily increased among adolescent girls yet remain lower in adolescent boys. In 2007, among girls aged 13-17 years, only 25.1% initiated the vaccine series and 5.9% completed the series, compared with 60.0% initiation and 39.7% completion in 2014 [5]. In 2014, vaccination rates among boys aged 13-17 years were 41.7% initiation and 21.6% completion [5].

A substantial body of communication research demonstrates that mediated communication reflects, but also serves to shape, popular understanding of important issues, including health [6-10]. As media evolves, more information about HPV vaccination is shifting to digital platforms on the Internet, in the form of websites, personal blogs, and social media. Of particular concern is the accuracy or viewpoint of the information. An analysis of the top search results about the HPV vaccine returned from Google, Yahoo, Bing, and Ask.com found that, while the majority of websites (57%) maintained a neutral tone about vaccination, a significant number of sites contained inaccuracies or mentioned conspiracy theories about the vaccine [11]. Users on Twitter, a popular microblogging social media platform, communicate about a range of topics, and there is strong evidence that communication includes a sizeable discourse on public health research and practice, including surveillance [12-16] and information dissemination [17-19].

Health information consumed on social media may be especially influential for segments of society such as younger populations, as well as ethnic and racial minorities, who may be less likely to access health information through formal news sources, health care providers, and other more traditional resources. For example, Latino, African American, and younger populations are more likely than white and older respondents to use mobile technologies for health information [20]. Furthermore, young users, along with minority users, disproportionately access Twitter on mobile devices. This is no surprise, as minority audiences are among the highest users of mobile technologies and social networking platforms [21,22].

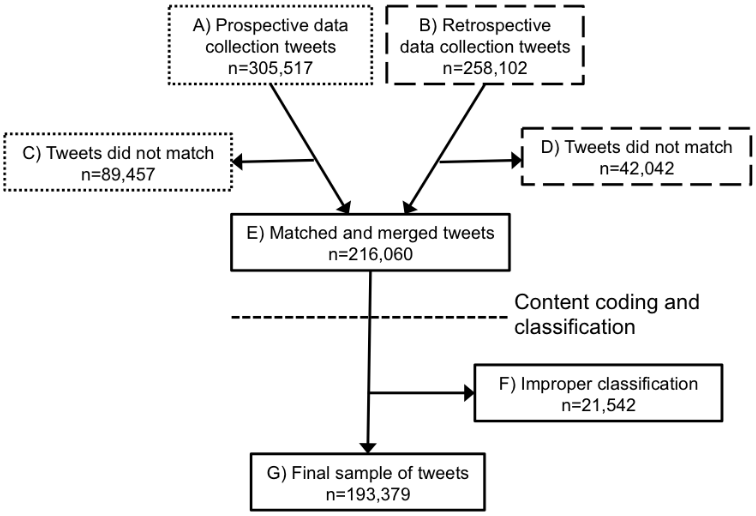
A major motivation for this study was the evolving nature of Web-based health information and the opportunity to better understand this area of inquiry through interdisciplinary research. That is, the ways in which Web-based health information is searched for and consumed are no longer limited to a 1-way or static process (ie, using a search engine). Rather, more and more Web information seekers are turning to more dynamic informational sources, including social media, blogs, and forums, to access but also respond to health information [23]. A recent nationwide survey conducted by the Pew Research Center, focusing on how adults in the United States use Web-based resources for health information, found that among Web-based health information seekers, 16% tried to find others who might share the same health concerns, 30% consulted Web-based reviews or rankings of health care services or treatments, and 26% read or watched someone else’s experience about a health issue [23]. Instinctively, public health researchers and practitioners are beginning to examine how health information is generated and disseminated via Twitter. There is a growing evidence base detailing methods for data collection and analysis using social media platforms, and in this study we sought to further this literature by describing a novel approach to data collection using two data collection tools.

These types of analyses and studies are particularly useful for the public health community, in tracking the dissemination of information about vaccination across populations and to gauge receptivity to vaccination messages. Researchers are just beginning to assess the extent and type of discourse about the HPV vaccine on Twitter, although methods are varied [24-27]. This emerging area of communication research provides an opportunity for interdisciplinary teams among the fields of public health, health communication, and data science to strengthen the science and methodology in this growing area of research. Thus, the purposes of this study were to quantify HPV vaccine communication on Twitter, specifically focusing on (1) sentiment, (2) side effects, and (3) prevention and protection, and to describe a novel methodology using two data collection methods to analyze Twitter data.

Methods

We used two methods to collect and validate Twitter data related to HPV vaccination, detailed in Figure 1. The first method used prospective data collection (Figure 1, box A) with a proprietary software program developed by Black and colleagues [28], and the second method employed retrospective data collection (Figure 1, box B) through Microsoft Research. To identify tweets related to HPV vaccination, we used the following keyword search terms and hashtags: “HPV,” “#HPV,” “HPV vaccine,” “#HPVvaccine,” “HPV shot,” “#hpvshot,” “Gardasil,” “#Gardasil,” “Cervarix,” and “#Cervarix.” We developed these keywords by drawing from previous research in content analyses of HPV print and Web-based news sources, balancing the general HPV-related topics with vaccine-specific information [29,30].

Figure 1. Flowchart detailing data collection, merging, and cleaning of final dataset of tweets related to human papillomavirus vaccination.



Prospective Data Collection

The proprietary software was developed to capture Twitter messages in real time, or prospectively, based on keyword search terms. Using the Twitter Search API (Twitter, Inc), the software collected publicly available data that contained tweets that were 6 or fewer days old, although it may include tweets up to 9 days old. The software captured not only the message posted on Twitter, including hashtags (#) and mentions (@), but also features of the message, such as information about the Twitter user; the date, location, and language of the tweet; and the number of times the message was retweeted. Information on the software architecture is detailed elsewhere [28]. As the software used the Twitter Search API, we used the software to conduct prospective data collection at least once a week during the study period (August 1, 2014 to July 31, 2015). We collected a total of 305,517 tweets using prospective data collection.

Retrospective Data Collection

In August 2015, after completing prospective data collection, Microsoft Research accessed Twitter Firehose to extract data using the same keyword search terms. Twitter Firehose allowed access to all tweets during the study timeline (August 1, 2014 to July 31, 2015), except those deleted by the users. We collected a total of 258,102 tweets using retrospective data collection.

Data Integration and Validation

All tweets contained a unique tweet identification (ID) number. Using this unique tweet ID, we merged and matched the prospective and retrospective datasets to produce a final dataset of n=216,060 tweets (Figure 1, box E). We included retweets in this final dataset as long as they had a unique tweet ID. The final dataset included 71.72% (89,457/305,517) of the prospective data and 83.71% of the retrospective data (42,042/258,102). Tweets that did not match were excluded from the final dataset (Figure 1, boxes C and D). Possible reasons for tweets not matching from the two data sources were that (1) user-selected language preferences changed between the prospective and retrospective collection periods, and thus the language variable did not match, (2) gaps longer than 7 days in prospective data collection due to study team scheduling would result in missing tweets no longer available through the Twitter Search API, and (3) some tweets had been deleted. Importantly, deleted tweets or tweets from suspended users that were captured in the prospective data collection were omitted from the final sample because they would not have been captured in the retrospective data collection (per Twitter’s user and data policy). We included English-language tweets, regardless of specific location, in the study; we removed tweets in all other languages from the final dataset.

Content Coding

Data captured included both the tweet content or message itself, contained within 140 characters, and information on the characteristics of the communication and sender. We developed a codebook to classify the content of the 140 characters, and in this study we report on sentiment toward the HPV vaccine (positive, negative, neutral, or no mention), side effects discussed, and prevention or protection discussed. We derived the coding system from previous content analysis research conducted by study team members about the HPV vaccine [29,30], although in print rather than in social media, as well as published Twitter content analysis research [13]. Table 1 details the sample codebook with features including variable description, along with tweet examples. “No mention” is included as part of sentiment because we wanted to situate vaccine sentiment within the larger HPV communication environment on Twitter. Therefore, we are able to quantify how much of the HPV discussion was vaccine focused and how this compares with HPV communication in general.

Table 1. Content classification codebook with feature description and tweet example for tweets related to human papillomavirus (HPV) vaccination.

Feature	Description	Example tweet
Sentiment		
Positive	The tweet contains supportive messages about the HPV vaccine and encourages its uptake	1. Not only does the HPV vaccine protect against human papillomavirus, but it also reduces the risk of cancers 2. #HPV vaccine can be #cancer prevention! Parents, #vaccinate your children at ages 11-12
Negative	The tweet contains disparaging messages about the HPV vaccine or discourages its uptake	1. Healthy 12-year-old girl dies shortly after receiving HPV vaccine 2. RT[a] @CBCHealth: The Gardasil Girls: How Toronto Star story on young women hurt public trust in vaccine http://t.co/...
Neutral	The tweet's text holds no subjective opinions about the vaccine—purely facts repeated from sources	1. State officials unveil campaign for HPV vaccination http://t.co/0I2sAWGXYs 2. RT @DrJenGunter: About 10% boys have received 3 doses HPV vax
No mention	The tweet does not mention the HPV vaccine	1. RT @Forbes: HPV is truly indiscriminate 2. RT @CDCSTD: #Women: get screened & talk w/ your friends about the link between #HPV & cervical #cancer
Side effects	The tweet refers to side effects caused by the HPV vaccine or effects that may be unknown to the user	1. Healthy 12-year-old girl dies shortly after receiving HPV vaccine 2. RT @ksbrowneyedgirl: It can happen to your child...to your family...#OneLess #Gardasil #CDCwhistleblower #vaccine...
Prevention/protection	The tweet refers to the extent to which the HPV vaccine will protect the user from or prevent negative health outcomes	1. Single HPV jab could prevent 70% of cervical cancers (http://t.co/Hg0KSlIk2A) 2. A new HPV vaccine prevents nine strains of the virus http://t.co/ZFGvVqlq0U

[a] RT: retweet.

Next, 5 members of the team independently coded the same random sample of tweets (n=50) for each of the codebook variables. Interrater reliability was high (>0.8 on each variable) among all 5 coders, indicating that the codebook is systematic and replicable. Based on developed coding procedures, 2 study members manually coded additional tweets for the purpose of developing classification models. This was an iterative process that involved manual coding of tweets, then computer-assisted coding, followed again by manual coding to refine the classification results. After each round of coding, we analyzed randomly selected tweets to validate model classifications. In total we coded 1470 tweets manually over 4 iterative rounds.

We used the 1470 manually coded tweets to develop a machine learning classifier for each variable in the codebook. Binary variables were classified using a linear classifier (Moore-Penrose pseudoinverse), while a decision tree was applied to variables with more than two categorical responses. Multimedia Appendix 1 details features of the classifiers.

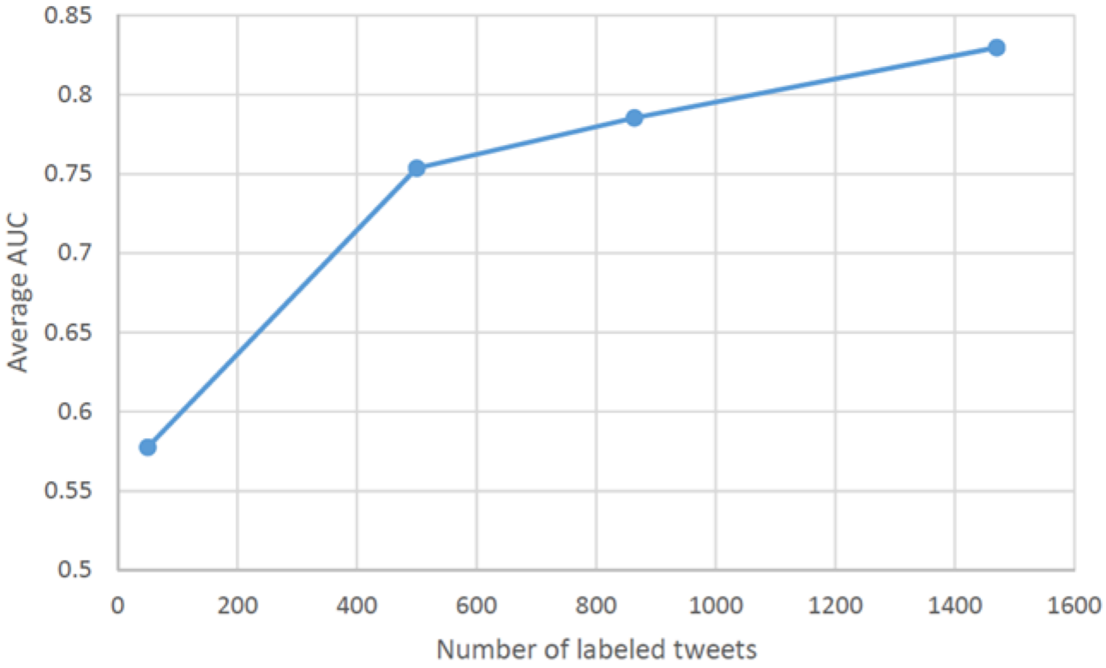
We evaluated the classifiers using leave-one-out estimation; that is, we trained classifiers on (n –1) samples and tested them on the remaining sample, repeating the process n times without replacement [31]. The accuracy of the classifications for binary variables was evaluated by the area under the receiver operating characteristic curve (AUC) and multiple-valued attributes via the fraction of errors. Figure 2 shows that the average AUC improved as a function of the number of manually coded tweets. In this study, we investigated 3 variables classified through this process: HPV vaccine sentiment (AUC=0.918), side effects (AUC=0.739), and prevention/protection (AUC=0.774).

After evaluating the performance of the machine classifiers, we then applied these to the full set of 216,060 tweets in the study, except for the 1470 manually labeled tweets. Based on the coding scheme, no study variables should have classification values rounding to zero; however, this was the case for some of values, indicating improper classification based on the coding scheme. Therefore, we randomly sampled 300 tweets with rounded classification values of zero and discerned no clear pattern within any of the unrounded classification values. That is, the tweets with values rounding to zero were not consistently supposed to be coded as 1 or another value. Due to this inconsistent classification pattern and to provide conservative estimates, we discarded all observations with a rounded classification of zero for sentiment, side effects, and prevention/protection (Figure 1, box F), leaving us with a final study sample of n=193,379 (Figure 1, box G).

Finally, for each tweet collected, we documented the words in the 140-character tweet text most associated with each keyword. We did this by computing the probability that a word would appear in tweets that contain the keyword, compared with the probability of that word in the entire corpus of tweets (n=193,379). To compute the probability, we counted the number of times each word appeared in the given set of tweets and divided this count by the total number of words in the set.

We used the statistical package Stata 14 (StataCorp LP) to analyze differences in tweet characteristics by tweet sentiment. To determine significant differences by sentiment, we used chi-square tests for the counts and categorical variables, analysis of variance for the continuous variables, and nonparametric equality of medians for the medians.

Figure 2. Average area under the receiver operating characteristic curve (AUC) as a function of manually coded tweets.



Results

A total of n=193,379 English-language tweets were collected, classified, and analyzed between August 1, 2014 and July 31, 2015. Figure 3 details the percentage of tweets that included the specified keyword search term. The keyword categories are not mutually exclusive, and the same tweet could be captured in multiple keyword searches. In our final dataset, the potential overlap of tweets across keyword searches was reconciled during the deduping and merging process. Over 88.64% (191,515/216,060) of the final dataset included the keyword search term HPV, and nearly 34.91% (75,433/216,060) included HPV vaccine. Based on the keyword query process, all tweets captured by the keyword HPV vaccine would by definition also be captured by the keyword HPV, as both searches share HPV.

Table 2 displays the words most associated with each keyword search term. Results show that associated words varied with each keyword, with HPV being associated with personal words such as *I*, *me*, and *have*, and #HPV being associated with *January* (cervical cancer awareness month), *prevent*, and *learn*. Words associated with the specific vaccine-related keywords (ie, HPV vaccine, Gardasil, and Cervarix) varied greatly, with more positive and preventive words being associated with HPV vaccine (ie, *read*, *to prevent*, *for girls*), and more negative words being associated with Gardasil and Cervarix (ie, *cdcwhistleblower*, *exposed*).

Table 2. Most significant terms associated with each keyword in tweets related to human papillomavirus (HPV) vaccine.

Keyword	Most significant terms
HPV	me, my, have HPV, I, I have, been, got, read, have to, like
#HPV	been, cervicalcancer, disease, January, to prevent, just, learn, it's, vaccine to, time, out
HPV vaccine	read, been, to prevent, vaccine to, I, out, for girls, age, the HPV, HPV vaccine
#HPVvaccine	HPVvaccine, to have, please, cancer is, teens, vaccine for, cervicalcancer, linked to, getting, safe
HPV shot	shot, have to, got, my, I have, me, go, to get, like, I
#HPVshot	12 year, 13 year, a new, about, about HPV, about the, active, after, against, against HPV
Gardasil	shot, Gardasil, 13 year, me, cdcwhistleblower, my, want, I, have to, got
#Gardasil	13 year, cdcwhistleblower, Gardasil, HPVvaccine, 100, need to, her, think, life, please
Cervarix	exposed, need to, many, medical, need, to be, cdcwhistleblower, health, how, research
#Cervarix	exposed, need to, medical, many, need, to be, cdcwhistleblower, health, research, link between

Table 3 describes overall sample totals, as well as sample totals by vaccine sentiment in the tweet (positive, negative, neutral, and no mention). Between August 2014 and July 2015 there were a total of 78,643 unique users who tweeted about HPV, with an average of 2.5 tweets per user. The average number of followers per user was 6569, compared with the median number of followers per user of 443. This large difference between the average and median number of followers indicates that, in our sample, the number of followers per user was heavily skewed to the right (toward a small number of highly followed users).

Moreover, the number of followers differed significantly by sentiment (*P*<.001), with a greater average number of followers exposed to positive sentiment than to negative sentiment (8022 vs 4772, respectively). A majority of tweets included a URL (138,059/193.379, 71.39%), nearly half included a hashtag (86,966/193.379, 44.97%), and just over half included a mention (112,049/193.379, 57.94%). Additionally, there were an average of 12 retweets per tweet and a median of 1 retweet per tweet. Tweets that did not mention the HPV vaccine had the highest average retweet count of 22 retweets per tweet.

Table 3. Characteristics based on metadata and classification analysis, n=193,379.

Characteristic	Tweet sentiment toward vaccine				Total	*P* value
	Positive	Negative	Neutral	No mention[a]		
HPV vaccine tweets						
Total tweets, n	75,393	48,940	25,110	43,936	193,379	<.001
Tweets with sentiment, %	38.99	25.31	12.98	22.72	—	
Users						
Total users, n	36,283	24,010	15,045	25,954	78,643[b]	<.001
Average HPV vaccine tweets per user	2.1	2.0	1.7	1.7	2.5	<.001
Average followers per user	8022	4772	6093	6352	6569	<.001
Median followers per user	459	467	445	381	443	<.001
Tweet contents						
Includes at least 1 link, n	57,800	34,491	18,898	26,870	138,059	<.001
Link, %	76.66	70.48	75.26	61.16	71.39	
Includes at least 1 hashtag, n	36,638	21,523	10,890	17,915	86,966	<.001
Hashtag, %	48.60	43.98	43.37	40.78	44.97	
Includes at least 1 mention, n	44,558	31,085	12,659	23,747	112,049	<.001
Mention, %	59.10	63.52	50.41	54.05	57.94	
Mentions prevention/protection, n	20,425	5647	6477	3209	35,758	<.001
Prevention/protection, %	27.09	11.54	25.79	7.30	18.49	
Mentions side effects/unknowns, n	12,921	22,726	3787	2619	42,053	<.001
Side effects/unknowns, %	17.14	46.44	15.08	5.96	21.75	
Retweets						
Average number of retweets per tweet	9.7	9.1	7.8	22.0	12.1	<.001
Median number of retweets per tweet	1	1	0	0	1	<.001

[a] No mention of HPV vaccine, but mention of HPV.

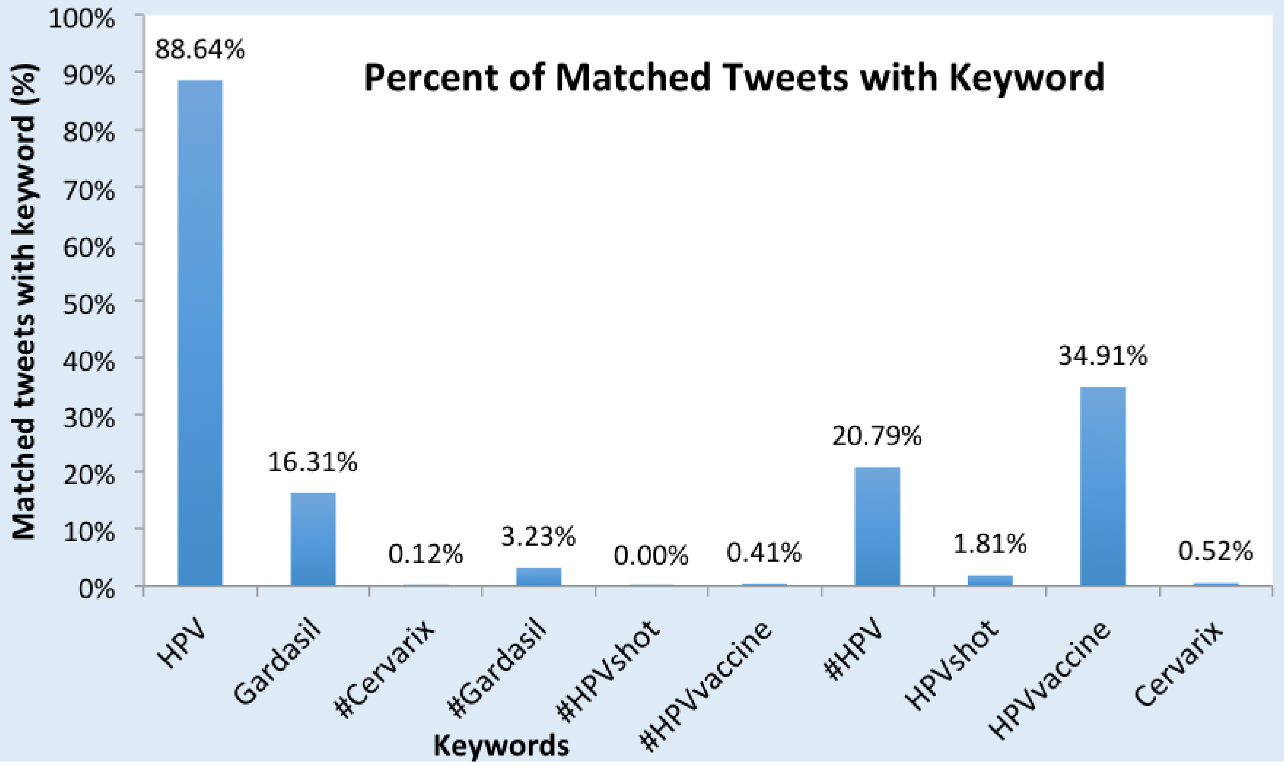
[b] Total differs from sum of totals because some users tweeted in multiple categories.

Table 3 also displays results by tweet sentiment. Positive, negative, or neutral sentiment describes how the HPV vaccine was communicated, and no mention indicates that the HPV vaccine *was not mentioned* (thought HPV *was mentioned*). Positive sentiment toward the vaccine was the largest type of sentiment in the sample, with 75,393 positive tweets (38.99% of the sample). Negative sentiment was the second largest type with 48,940 tweets (25.31% of the sample). Many more users participated in positive sentiment than in negative sentiment (36,283 vs 24,010 users, respectively).

Tweets coded as having positive sentiment toward HPV vaccine and no mention of HPV vaccine had a significantly higher use of URLs (57,800/75,393, 76.66%, and 18,898/25,110, 75.26% containing links, respectively) as compared with negative sentiment (34,491/48,940, 70.48%) (*P*<.001). The use of hashtags and mentions in tweets was fairly consistent across sentiment, with positive sentiment showing the greatest use of hashtags (36,638/75,393, 48.60%) and negative sentiment showing the greatest use of mentions (31,085/48,940, 63.52%).

When examining sentiment by tweet content, positive and neutral tweets had the largest percentage of mentions of prevention/protection (20,425/75,393, 27.09%, and 6477/25,110, 25.79%, respectively), compared with only 11.54% (5647/48,940) of negative tweets (*P*<.001). Nearly one-half (22,726/48,940, 46.44%) of negative tweets mentioned side effects, compared with only 17.14% (12,921/75,393) of positive tweets and 15.08% (3787/25,110) of neutral tweets.

Figure 3. Percentage of tweets in the final sample by keyword search.



Discussion

This study took place at the intersection of public health, health communication, and data science, and demonstrated the application of a novel methodology for collecting and integrating Twitter data from multiple sources, as well as supporting prior research showing the use of content and classification analysis. Building upon prior studies, we demonstrated the application of similar analyses and techniques across health domains [25,32], as well as expanding data collection and integration procedures. Using tweet characteristics, as well as content analysis based on classification models, we examined the potential reach and classified the nature of HPV vaccine-related tweets from August 2014 to July 2015 (n=193,379).

As our results show, words matter from both a data collection standpoint and a content perspective. The vast majority of our data contained the keyword HPV (88.64%), and thus future studies may be able to limit their keyword search to this single word, especially if resources are limited. Additionally, our findings show that different keywords are associated with different word clusters. HPV, for example, was associated with personal words such as *me*, *I*, and *have*, whereas #HPV was associated with more awareness-raising words, such as *January*, *prevent*, and *learn*. The hashtag (#) is an important feature on Twitter that categorizes tweets based on keyword and makes it easier to search other tweets with that same hashtag and keyword. By including #HPV in a tweet, users are able to click on #HPV and read other messages that have also included #HPV, thus acting as a “social search” function. This search feature may play an important role in raising awareness, as demonstrated by the associated word clusters. Importantly, brand-specific HPV vaccines were associated with more negative words on Twitter (ie, *cdcwhistleblower*, *exposed*), whereas the general keyword HPV vaccine was not. This could be important to understand when crafting messages about HPV and the vaccine: including brand-specific vaccines may encourage or lead to a more negative space in Twitter conversations.

Despite some research demonstrating that Web-based vaccine information can have a largely negative sentiment [11], our findings show that a great percentage of tweets about HPV vaccine had a positive sentiment, helping to validate findings on the same topic [24-27]. Furthermore, over one-quarter of these positive HPV vaccine tweets mentioned prevention or protection. In addition, positive tweets had, on average, many more followers than negative tweets, indicating the potential for a greater reach and more exposure of positive tweets than of negative tweets.

There is also an important relationship between tweet sentiment and tweet content: many more tweets that were classified as positive mentioned information about prevention or protection, whereas tweets classified as negative included a much greater discussion about side effects. This can be important information for health promotion and communication campaigns, specifically in terms of tailoring a message and joining a particular conversation. As tweets that contain information on side effects are more likely to be part of a negative conversation, tweeting “side effects are minimal,” that is, downplaying negative sentiment, may not be the most effective way to communicate this information. Importantly, for the “no mention” category (only discussion of HPV and not the vaccine), there was a very low percentage of tweets that mentioned prevention or protection as well as side effects. This is an important data analysis check in terms of the validity of our classification models, as we would not expect a large percentage of this content because it relates almost exclusively to the vaccine.

On Twitter, the distribution of followers by user is often highly skewed to the right; that is, few users (often dubbed the elite users, such as celebrities, organizations, and other high-profile users) have many more followers than the majority of Twitter users [33,34]. Our data support this distribution, as the average number of followers per user is much higher than the median number of follower per user (6569 compared to 443). When comparing average number of followers by sentiment, our findings show that tweets with positive sentiment toward the vaccine had a much higher average follower count than did tweets with negative sentiment, averaging 8022 followers compared to only 4772 followers, respectively. This suggests that more influential users, as measured by the number of followers, are tweeting about the vaccine more positively than negatively. This is an important finding, as number of followers is a proxy for reach, and thus more Twitter users are potentially being exposed to more positive sentiment than negative [33,34]. However, when examining sentiment distribution by the median, the numbers of followers are much closer, at 459 (positive) and 467 (negative). This suggests a more evenly distributed reach, or exposure, of positive and negative sentiment among typical Twitter users.

Social features on Twitter, including hashtags, mentions, and links, are important features to help with message dissemination and reach. Tweets classified as having positive sentiment accounted for the greatest percentage of tweets with links: 76.66% of positive tweets contained at least one link. Links are used to connect to other Web sources, often increasing the likelihood of interactivity (ie, shared by a retweet) and serving as a source information to support and corroborate the veracity of tweet content [35]. Tweet link content needs to be examined in order to examine the nature of links. Conversely, tweets classified as having negative sentiment had the greatest percentage of mentions per tweet: 63.52% (31,085/48,940) of negative tweets contained at least one mention. Mentions are a way to communicate directly to other users (more directed messaging and communication), and may serve as a way for agenda setters and opinion leaders to emerge in a network [36], bringing more attention to themselves and their message [37], and to control message diffusion [38]. Use of mentions may be a mechanism for negative and alternative messages about the HPV vaccine, which are in the minority, to appear to gain clout and recognition.

Measuring retweets is a way to quantify reach and dissemination [34,37,38], and in our sample “no mention” had the highest average number of retweets per tweet. “No mention” captures messages about HPV in general and not specifically related to HPV vaccine. This could potentially indicate an interest in disseminating information about the virus in general, or could be related to cervical cancer awareness and screening.

Understanding the content and implications of conversations that form around HPV vaccination on social media can aid health organizations and health-focused Twitter users in creating a meaningful exchange of ideas and having a significant impact on vaccine uptake. As HPV vaccination campaigns continue to use social media platforms, it is important to understand trends in social media communication, particularly across media platforms. In terms of public health surveillance, our study demonstrated that, despite an often negative-leaning frame and discussion of HPV vaccine on social media, the greatest percentage of HPV vaccine tweets are positive. Understanding effective dissemination channels will help connect campaigns with “elite” users and media who have many followers, and consequently may lead to a wider reach of message.

Limitations

While this study contributes to interdisciplinary research and methods, there are a few limitations worth noting. First, when merging tweets from the two data sources, we excluded 29.28% (89,457/305,517) of the prospective data and 16.29% (42,042/258,102) of the retrospective data due to nonmatching tweet IDs. To investigate why tweet IDs did not match and were thus excluded from the final sample, we examined excluded tweets from both data collection methods. We randomly selected 1000 tweets from the prospective data, which we identified through Twitter Search API, and found that many were excluded due to English-language misclassification (456/1000, 45.6%), deletion of the tweet (324/1000, 32.4%), no keyword being matched in the text (78/1000, 7.8%), and unexplained reasons (142/1000, 14.2%). We examined all tweets from the retrospective data, which were identified through Twitter Firehose, and found that many were excluded due to gaps in time of the prospective data (9821/42,042, 23.36%) and a majority for unexplained reasons (32,221/42,042, 76.64%).

While excluding tweets from the final sample may be a limitation, our study shows that the majority of tweets captured using the Twitter Search API, which is accessible to the public, was validated against the gold standard of Twitter Firehose. According to Twitter, Firehose contains “all public statuses,” compared with Twitter Search API, which only “offer samples of the public data flowing through Twitter.” Twitter Search API is further limited by the number of queries that can be made to it and the number of responses it returns. Therefore, for example, if there is a surge in the use of a keyword, only some of the tweets using it will be returned using Twitter Search API. For this reason, Firehose is considered the gold standard for Twitter data collection.

Second, when applying the classification models to our sample, we may have misclassified some of the tweets or not classified some at all. To limit misclassification, we used an iterative process that included multiple rounds of human coding to strengthen the computer classification models, and we reached on average 80% accuracy with each model. Tweets that were not classified with an adequate level of certainty (ie, above 70%) were excluded from the final sample. This process allowed for our entire sample of tweets to be classified as opposed to a randomly selected subsample.

Third, when discussing reach and exposure of health messages on Twitter, it is important to note that, when a user tweets, it does not mean that all of their followers will read the tweet. Thus, it is most accurate to refer to potential reach and exposure, as opposed to actual reach. In addition, analyzing the follower network of each user and tweet would provide additional information on reach; however, that was beyond the scope of this study.

Conclusion

Using and leveraging social media to detect health trends, as well as communicate important health information, is a growing area of research in public health. This area of research is inherently interdisciplinary, and this study supports this movement by applying public health, health communication, and data science approaches to extend methodologies across fields. Building on this particular study, future research will need to further examine how various stakeholders, including parents, youth, health care providers, and health care systems, communicate about the HPV vaccine and identify opportunities to strengthen vaccine uptake and completion. Furthermore, identifying how communication trends are associated with behavioral outcomes, that is, actual vaccine uptake, will be an important next phase of this area of inquiry.

Acknowledgments

This study was supported by Grant #IRG 14-251-07-IRG from the American Cancer Society.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Features of the classifiers.

[PDF File (Adobe PDF File), 19KB - jmir_v18i12e318_app1.pdf]

References

1. Gerend MA, Magloire ZF. Awareness, knowledge, and beliefs about human papillomavirus in a racially diverse sample of young adults. J Adolesc Health 2008 Mar;42(3):237-242. [doi: 10.1016/j.jadohealth.2007.08.022] [Medline: 18295131]
2. Markowitz LE, Dunne EF, Saraiya M, Chesson HW, Curtis CR, Gee J, Centers for Disease Control and Prevention (CDC). Human papillomavirus vaccination: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 2014 Aug 29;63(RR-05):1-30 [FREE Full text] [Medline: 25167164]
3. Centers for Disease Control and Prevention (CDC). Human papillomavirus vaccination coverage among adolescent girls, 2007-2012, and postlicensure vaccine safety monitoring, 2006-2013 - United States. MMWR Morb Mortal Wkly Rep 2013 Jul 26;62(29):591-595 [FREE Full text] [Medline: 23884346]
4. Petrosky E, Bocchini JA, Hariri S, Chesson H, Curtis CR, Saraiya M, Centers for Disease Control and Prevention (CDC). Use of 9-valent human papillomavirus (HPV) vaccine: updated HPV vaccination recommendations of the advisory committee on immunization practices. MMWR Morb Mortal Wkly Rep 2015 Mar 27;64(11):300-304 [FREE Full text] [Medline: 25811679]
5. Reagan-Steiner S, Yankey D, Jeyarajah J, Elam-Evans LD, Singleton JA, Curtis CR, et al. National, regional, state, and selected local area vaccination coverage among adolescents aged 13–17 years—United States, 2014. MMWR Morb Mortal Wkly Rep 2015 Jul 31;64(29):784-792 [FREE Full text] [Medline: 26225476]
6. Rubin AM, Haridakis P. Mass communication research at the dawn of the 21st century. In: Communication Yearbook 24. Thousand Oaks, CA: SAGE; 2001:73-97.
7. Snyder LB, Hamilton MA, Mitchell EW, Kiwanuka-Tondo J, Fleming-Milici F, Proctor D. A meta-analysis of the effect of mediated health communication campaigns on behavior change in the United States. J Health Commun 2004;9 Suppl 1:71-96. [doi: 10.1080/10810730490271548] [Medline: 14960405]
8. Love B, Himelboim I, Holton A, Stewart K. Twitter as a source of vaccination information: content drivers and what they are saying. Am J Infect Control 2013 Jun;41(6):568-570. [doi: 10.1016/j.ajic.2012.10.016] [Medline: 23726548]
9. Thackeray R, Burton SH, Giraud-Carrier C, Rollins S, Draper CR. Using Twitter for breast cancer prevention: an analysis of breast cancer awareness month. BMC Cancer 2013;13:508 [FREE Full text] [doi: 10.1186/1471-2407-13-508] [Medline: 24168075]
10. Myslín M, Zhu S, Chapman W, Conway M. Using Twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
11. Madden K, Nan X, Briones R, Waks L. Sorting through search results: a content analysis of HPV vaccine information online. Vaccine 2012 May 28;30(25):3741-3746. [doi: 10.1016/j.vaccine.2011.10.025] [Medline: 22019758]
12. Heaivilin N, Gerbert B, Page JE, Gibbs JL. Public health surveillance of dental pain via Twitter. J Dent Res 2011 Sep;90(9):1047-1051 [FREE Full text] [doi: 10.1177/0022034511415273] [Medline: 21768306]
13. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]

14. Prieto VM, Matos S, Álvarez M, Cacheda F, Oliveira JL. Twitter: a good place to detect health conditions. PLoS One 2014;9(1):e86191 [FREE Full text] [doi: 10.1371/journal.pone.0086191] [Medline: 24489699]
15. Ji X, Chun S, Geller J. Epidemic outbreak and spread detection system based on Twitter data. In: Proceedings of the HIS 2012. Berlin, Germany: Springer-Verlag; 2012 Presented at: HIS 1st International Conference on Health Information Science; April 8-10, 2012; Beijing, China p. 152-163. [doi: 10.1007/978-3-642-29361-0_19]
16. Gesualdo F, Stilo G, Agricola E, Gonfiantini MV, Pandolfi E, Velardi P, et al. Influenza-like illness surveillance on Twitter through automated learning of naïve language. PLoS One 2013;8(12):e82489 [FREE Full text] [doi: 10.1371/journal.pone.0082489] [Medline: 24324799]
17. Park H, Rodgers S, Stemmle J. Analyzing health organizations' use of Twitter for promoting health literacy. J Health Commun 2013;18(4):410-425. [doi: 10.1080/10810730.2012.727956] [Medline: 23294265]
18. Donelle L, Booth RG. Health tweets: an exploration of health promotion on twitter. Online J Issues Nurs 2012 Sep;17(3):4 [FREE Full text] [Medline: 23036060]
19. Harris JK, Mueller NL, Snider D, Haire-Joshu D. Local health department use of Twitter to disseminate diabetes information, United States. Prev Chronic Dis 2013;10:E70 [FREE Full text] [doi: 10.5888/pcd10.120215] [Medline: 23639765]
20. Fox S, Duggan M. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2012. Mobile health 2012 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2016-09-20] [WebCite Cache ID 6kffPoqjM]
21. Smith A. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013. Smartphone ownership 2013 update URL: http://www.pewinternet.org/files/old-media//Files/Reports/2013/PIP_Smartphone_adoption_2013_PDF.pdf [accessed 2016-09-20] [WebCite Cache ID 6kffYerha]
22. Duggan M, Smith A. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013. Social media update 2013 URL: http://www.pewinternet.org/files/2013/12/PIP_Social-Networking-2013.pdf [accessed 2016-09-20] [WebCite Cache ID 6kfgGk63X]
23. Fox S, Duggan M. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013. Health online 2013 URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2016-09-20] [WebCite Cache ID 6kfffQKEW]
24. Zhou X, Coiera E, Tsafnat G, Arachi D, Ong M, Dunn AG. Using social connection information to improve opinion mining: Identifying negative sentiment about HPV vaccines on Twitter. Stud Health Technol Inform 2015;216:761-765. [Medline: 26262154]
25. Surian D, Nguyen DQ, Kennedy G, Johnson M, Coiera E, Dunn AG. Characterizing Twitter discussions about HPV vaccines using topic modeling and community detection. J Med Internet Res 2016;18(8):e232 [FREE Full text] [doi: 10.2196/jmir.6045] [Medline: 27573910]
26. Dunn AG, Leask J, Zhou X, Mandl KD, Coiera E. Associations between exposure to and expression of negative opinions about human papillomavirus vaccines on social media: an observational study. J Med Internet Res 2015;17(6):e144 [FREE Full text] [doi: 10.2196/jmir.4343] [Medline: 26063290]
27. Mahoney LM, Tang T, Ji K, Ulrich-Schad J. The digital distribution of public health news surrounding the human papillomavirus vaccination: a longitudinal infodemiology study. JMIR Public Health Surveill 2015;1(1):e2 [FREE Full text] [doi: 10.2196/publichealth.3310] [Medline: 27227125]
28. Black A, Mascaro C, Gallagher M, Goggins SP. Twitter zombie: architecture for capturing, socially transforming and analyzing the Twittersphere. 2012 Presented at: 17th ACM International Conference on Supporting Group Work; October 27-31, 2012; Sanibel Island, FL, USA p. 229-238. [doi: 10.1145/2389176.2389211]
29. Casciotti DM, Smith KC, Tsui A, Klassen AC. Discussions of adolescent sexuality in news media coverage of the HPV vaccine. J Adolesc 2014 Feb;37(2):133-143. [doi: 10.1016/j.adolescence.2013.11.004] [Medline: 24439619]
30. Kelly BJ, Leader AE, Mittermaier DJ, Hornik RC, Cappella JN. The HPV vaccine and the media: how has the topic been covered and what are the effects on knowledge about the virus and cervical cancer? Patient Educ Couns 2009 Nov;77(2):308-313 [FREE Full text] [doi: 10.1016/j.pec.2009.03.018] [Medline: 19395221]
31. Duda RO, Hart PE, Stork DG. Pattern Classification. New York, NY: John Wiley & Sons; 2001.
32. Huang J, Kornfield R, Szczypka G, Emery SL. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control 2014 Jul;23 Suppl 3:iii26-iii30 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051551] [Medline: 24935894]
33. Wu S, Hofman JM, Mason WA, Watts DJ. Who says what to whom on Twitter. In: Proceedings of the 20th International Conference Companion on World Wide Web. New York, NY: ACM Press; 2011 Presented at: WWW '11; March 28-April 1; Hyderabad, India p. 705-714 URL: http://www.wwwconference.org/proceedings/www2011/proceedings/p705.pdf
34. Bakshy E, Hofman JM, Mason WA, Watts DJ. Everyone's an influencer: quantifying influence on Twitter. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. New York, NY: ACM; 2011 Presented at: WSDM '11; February 9-12, 2011; Hong Kong, China p. 65-74.
35. Suh B, Hong L, Pirolli P, Chi EH. Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In: Social Computing (SocialCom). Los Alamitos, CA: IEEE; 2010 Presented at: 2010 IEEE Second International Conference; August 20-22, 2010; Minneapolis, MN, USA p. 177-184. [doi: 10.1109/SocialCom.2010.33]

36. Meraz S, Papacharissi Z. Networked gatekeeping and networked framing on #Egypt. Int J Press/Politics 2013 Jan 27;18(2):138-166. [doi: 10.1177/1940161212474472]
37. Cha M, Haddadi H, Benevenuto F, Gummadi P. Measuring user influence in Twitter: the million follower fallacy. 2010 Presented at: 4th International AAAI Conference on Weblogs and Social Media (ICWSM); May 23-26, 2010; Washington, DC, USA p. 1-8 URL: http://snap.stanford.edu/class/cs224w-readings/cha10influence.pdf
38. Bastos MT, Raimundo RLG, Travitzki R. Gatekeeping Twitter: message diffusion in political hashtags. Media Cult Soc 2013 Mar 14;35(2):260-270. [doi: 10.1177/0163443712467594]

Abbreviations

AUC: area under the receiver operating characteristic curve
HPV: human papillomavirus
ID: identification

Edited by G Eysenbach; submitted 20.09.16; peer-reviewed by C Zhang, D Surian; comments to author 06.10.16; revised version received 04.11.16; accepted 23.11.16; published 05.12.16.

Please cite as:
Massey PM, Leader A, Yom-Tov E, Budenz A, Fisher K, Klassen AC
Applying Multiple Data Collection Tools to Quantify Human Papillomavirus Vaccine Communication on Twitter
J Med Internet Res 2016;18(12):e318
URL: http://www.jmir.org/2016/12/e318/
doi:10.2196/jmir.6670
PMID:27919863

Original Paper

The Association Between Internet Use and Ambulatory Care-Seeking Behaviors in Taiwan: A Cross-Sectional Study

Ronan Wenhan Hsieh[1*], MD; Likwang Chen[2*], PhD; Tsung-Fu Chen[3], SJD; Jyh-Chong Liang[4], PhD; Tzu-Bin Lin[5], PhD; Yen-Yuan Chen[6], MPH, MD, PhD; Chin-Chung Tsai[4], PhD

[1]Department of Medicine, National Taiwan University College of Medicine, Taipei, Taiwan
[2]Institute of Population Health Sciences, National Health Research Institutes, Miao-Li, Taiwan
[3]National Taiwan University College of Law, Taipei, Taiwan
[4]Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, Taipei, Taiwan
[5]Department of Education, National Taiwan Normal University, Taipei, Taiwan
[6]Graduate Institute of Medical Education & Bioethics, National Taiwan University College of Medicine, Taipei, Taiwan
[*]these authors contributed equally

Corresponding Author:
Yen-Yuan Chen, MPH, MD, PhD
Graduate Institute of Medical Education & Bioethics
National Taiwan University College of Medicine
#1, Rd. Ren-Ai sec. 1
Taipei, 100
Taiwan
Phone: 886 2 23123456 ext 63911
Fax: 886 2 23224793
Email: chen.yenyuan@gmail.com

Abstract

Background: Compared with the traditional ways of gaining health-related information from newspapers, magazines, radio, and television, the Internet is inexpensive, accessible, and conveys diverse opinions. Several studies on how increasing Internet use affected outpatient clinic visits were inconclusive.

Objective: The objective of this study was to examine the role of Internet use on ambulatory care-seeking behaviors as indicated by the number of outpatient clinic visits after adjusting for confounding variables.

Methods: We conducted this study using a sample randomly selected from the general population in Taiwan. To handle the missing data, we built a multivariate logistic regression model for propensity score matching using age and sex as the independent variables. The questionnaires with no missing data were then included in a multivariate linear regression model for examining the association between Internet use and outpatient clinic visits.

Results: We included a sample of 293 participants who answered the questionnaire with no missing data in the multivariate linear regression model. We found that Internet use was significantly associated with more outpatient clinic visits (P=.04). The participants with chronic diseases tended to make more outpatient clinic visits (P<.01).

Conclusions: The inconsistent quality of health-related information obtained from the Internet may be associated with patients' increasing need for interpreting and discussing the information with health care professionals, thus resulting in an increasing number of outpatient clinic visits. In addition, the media literacy of Web-based health-related information seekers may also affect their ambulatory care-seeking behaviors, such as outpatient clinic visits.

(J Med Internet Res 2016;18(12):e319) doi:10.2196/jmir.5498

KEYWORDS

mass media; Internet; literacy; outpatient clinic; ambulatory care

Introduction

The way people search for health information is constantly changing. Not long ago, people relied almost solely on physicians' advice to address patients' medical needs because medicine was a highly specialized field of knowledge, inaccessible and incomprehensible to the general population. However, recent studies have indicated that the picture is gradually changing [1].

According to an annual report published by the Taiwan Network Information Center in 2012, 15.94 million Taiwanese used the Internet, accounting for 77.25% of the Taiwanese population aged 12 years and older [2]. Among those Internet users, many had used the Internet for searching health-related information. Research reported that, in the United States in 2009, 74% of adults used the Internet, and 61% had searched for health-related information online. In particular, 49% had accessed a website for understanding a specific medical condition or problem [3]. As reported by the Pew Research Center in 2013, approximately 59% of Americans had searched online health information in the preceding year [4]. The Internet offers a wide variety of health-related information, allowing the general population to reaffirm what they have heard from their doctors and to make a thoroughly informed decision [5-7].

Compared with the traditional ways of gaining health-related information, such as visiting an outpatient clinic, reading newspapers or magazines, listening to radio or watching television shows, getting advice from neighbors or community or family members, the Internet is inexpensive, accessible, and conveys diverse opinions. On the other hand, more and more physicians use the convenience of the Internet to search drug information [8] and Web-based evidence sources such as UpToDate [9] for providing better medical care to patients.

Relevant studies on how increasing Internet use affects outpatient clinic visits have been inconclusive. Azocar et al [10] in 2003 concluded that patients' use of a behavioral health website significantly motivated them to increase their use of health care services. In 2008, Lee [11] reported that increasing use of the Internet was correlated with increasing outpatient clinic visits, although only age and sex were controlled in the data analysis. Another study analyzed the association between Internet use and the Internet users' health-seeking behaviors for those who searched the Web for different purposes. The authors reported that, for every 1-hour increment of searching the Internet, the likelihood for patients to visit physicians was increased by about 10% [12].

In contrast, some studies found a negative or no effect of using the Internet on the number of outpatient clinic visits. A survey conducted in the United States showed that 94% of the participants said their Internet use did not change the number of outpatient clinic visits they made [13]. Another survey conducted in Japan reported that 88.9% of the respondents thought there was no association between their Internet use and their frequency of visiting outpatient clinics or making phone calls to their physicians to inquire about health-related issues [14].

Given that the association between Internet use and health-seeking behaviors is still controversial, and that prior studies were conducted without sufficiently adjusting for potential confounding variables, we conducted this survey in which we derived the data from a sample randomly selected from the general population in Taiwan, with adjustment for most of the potential confounding variables. The aim of this study was to examine the role of Internet use in health-seeking behaviors as indicated by the number of outpatient clinic visits.

Methods

We derived the data for this study from the sixth cycle's second year (2011) survey of the research project Taiwan Social Change Survey (TSCS) [15]. The project was conducted by the Institute of Sociology, Academia Sinica, and sponsored by the Ministry of Science and Technology (formerly known as the National Science Council) in Taiwan. Each year, 2 modules were used for the questionnaires. For example, in 2009, the 2 modules were Social Inequality and Religion. A total of 2026 Social Inequality questionnaires and 1924 Religion questionnaires were finally answered by the participants randomly selected from Taiwan's general population.

The 2011 TSCS survey contained 2 modules: (1) the Family questionnaire, and (2) the Health questionnaire. Participants in the 2011 TSCS survey were randomly selected from the general population in Taiwan. Face-to-face structured interviews were conducted for all selected participants. In addition to shared questions, some of the survey questions in the questionnaire had been separated into 2 sets: set-A questions and set-B questions. Each set-A question has a similar set-B question. For example, question D1a1 in set A and question D1b1 in set B both inquire about participants' smoking status.

This was a cross-sectional study. The Health questionnaires collected for this study were answered by the participants who were randomly selected from Taiwan's general population. Participants in this study were numbered: the odd-numbered participants were assigned to answer set-A questions and the shared questions; and the even-numbered participants were assigned to answer set-B questions and the shared questions in the Health questionnaires. Our study used a secondary dataset collected in 2011 by the Institute of Sociology in Academia Sinica.

We selected the following variables: (1) reported confounding variables such as age [16,17], sex [13,16], annual income [16], educational level [14], self-reported health status, and chronic disease [17,18]; (2) background information of a participant such as residence and marital status; (3) social support variables such as the total number of family members and self-perceived neighbor support; and (4) attitudes toward the health care system.

For balancing the variables between the group of questionnaires without missing data and the group of questionnaires with missing data, we built a model for propensity score matching using multivariate logistic regression. Age and sex were the independent variables, and whether a questionnaire answered by a participant had missing data (the uncompleted group) or not (the completed group) was the dependent variable in the multivariate logistic regression model for propensity score matching. We obtained each participant's propensity score of being assigned to the completed group based on the multivariate logistic regression model. A participant in the completed group was matched to a participant in the uncompleted group using 1-to-1 nearest-neighbor matching without replacement and a caliper of 0.18 of the pooled standard deviation of the logit of the propensity scores. We examined whether the propensity score model had good discrimination by using the area under the receiver operating characteristic curve. We expected age and sex in the propensity score model to be balanced between the completed and uncompleted groups. The questionnaires in the completed group that were matched to the questionnaires in the uncompleted group were retained for further analysis.

For examining the linear association between an independent variable and the outcome variable, we calculated the Pearson correlation coefficient or Spearman rank correlation coefficient depending on the scale of an independent variable. The outpatient clinic visit was the outcome variable. We coded the outpatient clinic visit using a Likert scale ranging from 1 to 5, indicating the frequency of outpatient clinic visits from "never" to "several times a month," respectively. An independent variable with a *P* value of the correlation coefficient <.30 was eligible to enter the multivariate linear regression model. In addition, we examined the collinearity between 2 independent variables using the Spearman rank correlation coefficient: "How is your health status?" and "Do you have chronic diseases?" Only one of them stayed in the model if a significant collinearity was identified.

We conducted multivariate linear regression analysis for examining the association between Internet use and outpatient clinic visits, including the confounding variables selected by linearity and collinearity checks. We regarded *P*≤.05 as statistically significant. We conducted all statistical analyses in this study using STATA/MP 11.0 (StataCorp LP) for Windows PC. This study was approved by the Research Ethics Committee of National Taiwan University Hospital (201510102W).

Results

A total of 2199 participants joined this study to answer the Health questionnaires. Among the 2199 Health questionnaires, 1064 (48.39%) were even numbered and assigned to answer the set-B questions and the shared questions. Given that set-A and set-B questions were established using different scales—that is, a scale from 1 to 5 was established for set-A questions but a scale from 1 to 6 was established for set-B questions—we thus excluded the participants who answered set-B questions from further analysis. Among the 1135 participants who were odd numbered and assigned to answer the set-A questions and the shared questions on the Health questionnaires, 23 (2.03%) did not provide their age and were thereby excluded. Therefore, we finally included 1112 Health questionnaires for propensity score matching using multivariate logistic regression (Figure 1).

A total of 1112 Health questionnaires were eligible for this study: 555 questionnaires did not have missing data (the completed group); 557 questionnaires (50.09%) had missing data (the uncompleted group). The participants in the completed group were significantly younger than those in the uncompleted group (*P*<.001). More male participants were in the completed group than in the uncompleted group with a borderline significance (*P*=.06). To avoid the potential risk of selection bias, we conducted propensity score matching to balance age and sex between the 555 questionnaires in the completed group and the 557 questionnaires in the uncompleted group (Table 1).

We established a propensity score model for matching a questionnaire answered by a participant in the completed group (n=555) to a questionnaire answered by a participant in the uncompleted group (n=557). The propensity score model for the completed group and uncompleted group included 2 independent variables: age and sex. It showed good discrimination (area under the receiver operating characteristic curve=0.83). We identified 293 matched pairs: 293 respondents from the completed group and 293 respondents from the uncompleted group. The mean (SD) propensity scores before matching were 0.67 (0.22) for the completed group and 0.33 (0.25) for the uncompleted group (*P*<.001). The mean (SD) propensity scores after matching were 0.55 (0.22) for the 293 participants from the completed and 0.51 (0.22) for the 293 participants from the uncompleted groups (*P*=.07). Age and sex were not significantly different between the 293 participants from the completed group and the 293 participants from the uncompleted group. The 293 participants from the completed group were eligible for bivariate analysis and multivariate linear regression analysis.

In examining the linear relationship between each independent variable and the number of outpatient clinic visits (Table 2), we found that participants who rated themselves healthier were negatively associated with making more outpatient clinic visits (*P*<.003). In comparison, people with chronic diseases, such as diabetes, hypertension, or hyperlipidemia, tended to have made more outpatient clinic visits (*P*<.001). The collinearity check using the Spearman rank correlation coefficient for "What is your health status?" and "Do you have chronic diseases?" showed a correlation coefficient of –.18 (*P*=.002). We therefore excluded "What is your health status?" from further analysis.

Figure 1. Participant selection.

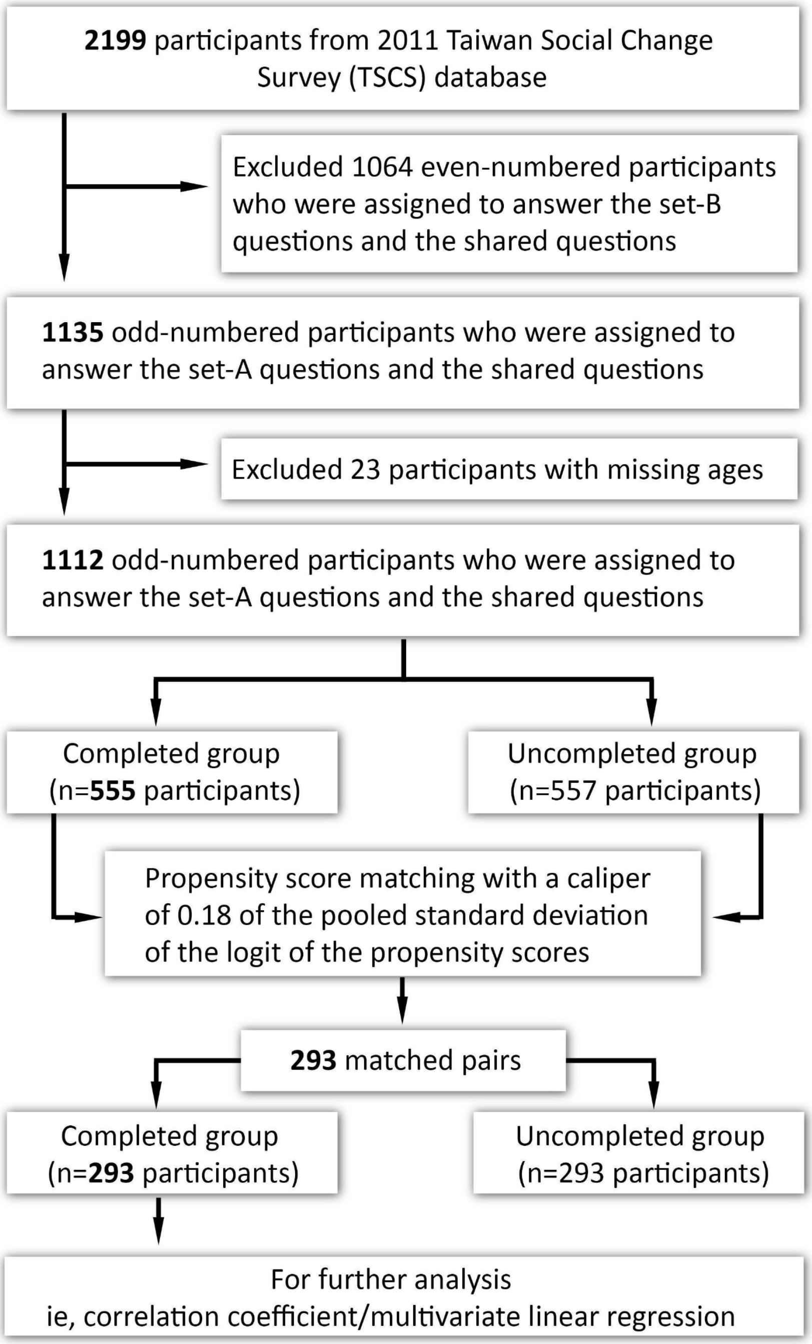


Table 1. Characteristics before and after propensity score matching for the completed group and uncompleted group.

Questionnaire item (score)	Before propensity score matching				After propensity score matching			
	Completed (n=555) Mean (SD) or n (%)	Uncompleted (n=557) Mean (SD) or n (%)	Missing data[a]	*P* value	Completed (n=293) Mean (SD) or n (%)	Uncompleted (n=293) Mean (SD) or n (%)	Missing data[a]	*P* value
How often have you visited a doctor in the past 12 months? n (%)			70	<.001			56	.47
Never (1)	27 (4.86)	13 (2.76)			16 (5.46)	8 (3.38)		
Once a year (2)	51 (9.19)	35 (7.19)			21 (7.17)	25 (10.55)		
Several times a year (3)	383 (69.01)	295 (60.57)			205 (69.97)	160 (67.51)		
Once a month (4)	78 (14.05)	118 (24.23)			42 (14.33)	34 (14.35)		
Several times a month (5)	16 (2.88)	26 (5.34)			9 (3.07)	10 (4.22)		
How many hours do you use the Internet every day? mean (SD)								
	2.90 (2.69)	2.87 (3.15)	473	.93	2.28 (2.39)	2.95 (3.18)	212	.04
How many hours do you watch TV news every day? mean (SD)								
	0.98 (0.96)	1.10 (1.07)	3	.045	1.08 (1.11)	1.14 (1.05)	2	.49
How many years of school education did you receive? mean (SD)								
	14.23 (2.89)	9.66 (3.79)	71	<.001	13.79 (3.16)	10.78 (3.54)	7	<.001
Age in years, mean (SD)								
	37.21 (13.03)	57.91 (16.58)	0	<.001	43.89 (13.17)	45.94 (12.35)	0	.06
Sex, n (%)			0	.06			0	.16
Female	265 (47.75)	297 (53.32)			163 (55.63)	146 (49.83)		
Male	290 (52.25)	260 (46.68)			130 (44.37)	147 (50.17)		
Residence, n (%)			2	<.001			2	.03
Rural	18 (3.24)	54 (9.73)			11 (3.75)	23 (7.90)		
Urban	537 (96.76)	501 (90.27)			282 (96.25)	268 (92.10)		
What is your marital status? n (%)			2	<.001			2	.008
Unmarried	245 (44.14)	59 (10.63)			77 (26.28)	57 (19.59)		
Divorced	17 (3.06)	27 (4.86)			11 (3.75)	23 (7.90)		
Widowed	2 (0.36)	88 (15.86)			2 (0.68)	10 (3.44)		
Married	285 (51.35)	372 (67.03)			201 (68.60)	196 (67.35)		
Other	6 (1.08)	9 (1.62)			2 (0.68)	5 (1.72)		
How many members, including you, are there in your family? mean (SD)								
	4.25 (1.86)	4.14 (2.13)	3	.35	4.20 (1.76)	4.23 (1.91)	3	.86
When you need help, your neighbors are willing to give you a hand. mean (SD)[b]								
	3.79 (0.89)	3.88 (0.84)	26	.09	3.84 (0.83)	3.82 (0.84)	21	.82
Are you satisfied with Taiwan’s health care system? mean (SD)[c]								
	3.46 (1.08)	3.62 (1.01)	13	.02	3.42 (1.12)	3.41 (1.05)	5	.91
Generally speaking, Taiwan’s physicians are trustworthy. mean (SD)[b]								
	3.56 (0.90)	3.49 (0.94)	9	.21	3.52 (0.90)	3.37 (1.01)	4	.07
Do you smoke? n (%)			2	.26			0	<.001
No	454 (81.80)	439 (79.10)			248 (84.64)	210 (71.67)		
Yes	101 (18.20)	116 (20.90)			45 (15.36)	83 (28.33)		

Questionnaire item (score)	Before propensity score matching				After propensity score matching			
	Completed (n=555) Mean (SD) or n (%)	Uncompleted (n=557) Mean (SD) or n (%)	Missing data[a]	*P* value	Completed (n=293) Mean (SD) or n (%)	Uncompleted (n=293) Mean (SD) or n (%)	Missing data[a]	*P* value
Do you drink alcohol? n (%)			3	<.001			0	.56
No	272 (49.01)	345 (62.27)			162 (55.29)	155 (52.90)		
Yes	283 (50.99)	209 (37.73)			131 (44.71)	138 (47.10)		
Do you chew betel nut? n (%)			0	<.001			0	<.001
No	515 (92.79)	475 (85.28)			277 (94.54)	238 (81.23)		
Yes	40 (7.21)	82 (14.72)			16 (5.46)	55 (18.77)		
Did you receive a self-paid health checkup in the past 3 years? n (%)			4	<.001			2	.30
No	403 (72.61)	458 (82.82)			222 (75.77)	231 (79.38)		
Yes	152 (27.39)	95 (17.18)			71 (24.23)	60 (20.62)		
What is your health status? mean (SD)[d]								
	2.63 (0.97)	2.44 (1.03)	2	<.002	2.61 (0.96)	2.58 (1.04)	0	.77
Do you have chronic diseases? n (%)			10	<.001			9	.25
No	437 (78.74)	294 (53.75)			215 (73.38)	196 (69.01)		
Yes	118 (21.26)	253 (46.25)			78 (26.62)	88 (30.99)		
In the past 12 months, did you seek medical assistance from complementary and alternative medicine? n (%)			61	.10			49	.61
No	343 (61.80)	331 (66.73)			181 (61.77)	156 (63.93)		
Yes	212 (38.20)	165 (33.27)			112 (38.23)	88 (36.07)		
Are you satisfied with your quality of life? n (%)			10	.92			3	.02
Very dissatisfied or dissatisfied	115 (20.72)	112 (20.48)			51 (17.41)	73 (25.17)		
Very satisfied or satisfied	440 (79.28)	435 (79.52)			242 (82.59)	217 (74.83)		
Is your household income higher than, lower than, or similar to other households in Taiwan? mean (SD)[e]								
	2.88 (0.63)	2.59 (0.75)	12	<.001	2.87 (0.64)	2.67 (0.71)	5	<.001

[a] The sample size of missing data in the uncompleted group next to itself for the item.

[b] Possible responses and their scores were strongly disagree (1), disagree (2), neutral (3), agree (4), and strongly agree (5).

[c] Possible responses and their scores were very dissatisfied (1), dissatisfied (2), neutral (3), satisfied (4), and very satisfied (5).

[d] Possible responses and their scores were bad (1), fair (2), good (3), very good (4), and excellent (5).

[e] Possible responses and their scores were much lower (1), lower (2), similar (3), higher (4), and much higher (5).

Table 2. Correlation coefficients between each independent variable and the number of outpatient clinic visits.

Variable	Correlation coefficient	*P* value
How many hours do you use the Internet every day?	.06	.27
How many hours do you watch TV every day?	.04	.48
How many years of school education did you receive?	–.10	.08
Age (1-year increment)	.11	.06
Sex (0=female, 1=male)	–.06	.34
Residence (0=rural, 1=urban)	–.01	.83
What is your marital status? (reference group: married)	–.04	.49
How many members, including you, are there in your family?	–.11	.06
When you need help, your neighbors are willing to give you a hand.	.02	.77
Are you satisfied with Taiwan’s health care system?	.01	.83
Generally speaking, Taiwan’s physicians are trustworthy.	.08	.15
Do you smoke? (0=no, 1=yes)	–.09	.11
Do you drink alcohol? (0=no, 1=yes)	–.07	.22
Do you chew betel nut? (0=no, 1=yes)	<.01	.99
Do you receive self-paid health checkup in the past 3 years? (0=No, 1=Yes)	.07	.22
What is your health status? (reference group: Bad)	–.18	.003
Do you have chronic diseases? (0=no, 1=yes)	.25	<.001
In the past 12 months, did you seek medical assistance from complementary and alternative medicine? (0=no, 1=yes)	.01	.83
Are you satisfied with your quality of life? (0=very dissatisfied/dissatisfied, 1=very satisfied/satisfied)	.06	.31
Is your household income higher than, lower than, or similar to other households in Taiwan?	.07	.24

After controlling for other confounding variables using multivariate linear regression for the 293 participants from the completed group, we found that Internet use was significantly associated with more outpatient clinic visits (*P*=.04) (Table 3).

The adjusted R^2 was .1195, indicating that 11.95% of the variance could be accounted for by this multivariate linear regression model.

Table 3. Multivariate linear regression on number of outpatient office visits (n=293)[a].

Variable	Coefficient	*P* value
How many hours do you use the Internet every day?	0.04	.04
How many years of school education did you receive?	–0.03	.03
Age (1-year increment)	0.01	.20
Sex (0=female, 1=male)	–0.09	.32
What is your marital status? (reference group: married)		
Unmarried	0.24	.08
Divorced	0.15	.50
Widowed	0.75	.15
Other	–1.22	.02
How many members, including you, are there in your family?	–0.02	.50
Generally speaking, Taiwan’s physicians are trustworthy.	0.01	.76
Do you smoke? (0=no, 1=yes)	–0.19	.12
Do you drink alcohol? (0=no, 1=yes)	–0.11	.21
Did you receive a self-paid health checkup in the past 3 years? (0=no, 1=yes)	0.12	.22
Do you have chronic diseases? (0=no, 1=yes)	0.43	<.001
Is your household income higher than, lower than, or similar to other households in Taiwan?	0.04	.51

[a] The value of adjusted R^2 for this multivariate linear regression model is .1195.

Discussion

Principal Findings

After controlling for age and sex using propensity core matching for the completed group and uncompleted group, and for other confounding variables using multivariate linear regression analysis, we found that spending more time using the Internet was significantly associated with making more outpatient clinic visits.

Health-Related Information in the Media

Media such as newspapers, magazines, journals, television, and radio report not only information on daily life and their audiences’ interests, but also health-related information. One of the most important ways the media may help patients to correctly interpret health-related information is to present the information in an unbiased manner [19,20]. Nevertheless, this is not usually the case. Web-based health information may not be as correct as the information shown in textbooks or academic journals [21].

Diem et al [22] reported that, in television programs, the survival rates of people receiving cardiopulmonary resuscitation were significantly higher than the most optimistic survival rates reported in the literature. Chen et al [23] also reported that information reported in the major newspapers in Taiwan regarding the use of life supporting treatments for patients who are critically ill is too optimistic as indicated by the probability of survival. In addition to health-related information regarding life supporting treatments in the media, Moynihan et al [24] studied media coverage of the benefits and risks for three medications in leading national newspapers, local newspapers, and television networks in the United States. They concluded that media stories about medications included inadequate or insufficient information about the benefits, risks, and costs of the drugs.

One theme emerged from the above studies [22-24], which is that the media tends to show biased health-related information. Accordingly, the audience’s ability to analyze and evaluate the messages shown on a wide variety of media is critical for telling whether health-related information in the media is biased or unbiased.

Internet Use and Outpatient Clinic Visits

The media plays an important role in raising awareness about health care services to patients and in shaping laypeople’s perceptions of and decision making about health care [25]. Our study found that laypeople who spend more time using the Internet are more likely to visit outpatient clinics. Several reasons may account for this phenomenon:

First, some studies reported that people who used the Internet to search for health-related information needed more help from health care professionals with interpreting and understanding the health-related information they obtained [11,26]. Another study reported that the controversial health-related information reported in the media might affect patients’ perceptions of and decision making about medical care [27]. In addition, several studies have shown that health-related information on the Internet is less likely to be accredited by the Health on the Net Foundation, and is therefore less reliable [28]. Such inconsistent quality may bring about significant anxiety for patients. As a result, patients who spend more time using the Internet may be more likely to visit outpatient clinics for clarification and interpretation of a kaleidoscope of health-related information. This, therefore, may increase physicians’ workloads, as they have to spend additional time discussing health-related information and reassuring patients [27].

Second, compared with nonseekers, adults seeking health-related information on the Web were more likely to rate themselves as having poor health [13]. Therefore, laypeople who spend more time using the Internet are more likely to visit outpatient clinics, not because of their Internet use, but because they tend to rate themselves as having poor health and, therefore, tend to seek professional advice about their health status.

Third, media literacy may account for our study result. Laypeople with better media literacy may not simply accept health-related information shown in the media. Instead, they may prefer to carefully digest health-related information obtained from the media by studying academic journals, seeking professional guidance by consulting health care professionals, and so on. Visiting outpatient clinics to obtain professional advice for health-related information shown in the media is the most convenient and least time-consuming way to get that guidance. As a result, people who frequently seek health-related information on the Web are significantly associated with making more outpatient clinic visits.

Strengths and Limitations

We conducted this study, based on a sample derived from a random sample of the general population in Taiwan, to examine the relationship between Internet use and health-seeking behaviors as indicated by the number of outpatient clinic visits. We used sophisticated statistical methods to minimize the threat to external validity due to missing data and to control a large variety of confounding variables. In addition, the study results extended what is already known and from previously reported academic work by providing new data and by controlling for some confounding variables that were not controlled for (ie, the total number of family members and whether physicians are perceived as trustworthy) in previous studies [11-13]. Accordingly, we are confident that the results of this study are convincing and generalizable.

Nevertheless, there are some limitations in this study. This was a cross-sectional study using a questionnaire for the participants to recall the frequency of their Internet use and the frequency of their outpatient clinic visits. The causal relationship between Internet use and outpatient clinic visits is not as strong as in the design of a prospective cohort study. In addition, the frequency of Internet use recalled by the study respondents did not necessarily indicate that they had been searching for health-related information. This potential inaccuracy should be further considered when applying the results of this study.

Second, recall bias may also affect the outcome of this study. It could simply occur due to the differences in the accuracy or completeness of participants' answers to the survey questions regarding their Internet use and outpatient clinic visits from the past year. Or, if participants were so sick that they had visited outpatient clinics several times, it would not have been easy for them to recall the frequency of their visits in the past year. Similarly, if participants only sometimes used the Internet, they might not have been able to recall the frequency of their Internet use.

Third, there may be concerns about the relationship between Internet use and the number of outpatient clinic visits, which was nonsignificant in univariate analysis (P=.27) but significant in multivariate analysis (P=.04). Several reasons may cause this phenomenon [29]: (1) interaction: we have checked the interaction between the time of using the Internet every day in hours and chronic diseases status, and identified that there is no interaction between these 2 variables (P=.57); (2) the effect of unbalanced sample size: the dataset for multivariate linear regression did not have unbalanced sample size; and (3) the influence of missing data: there were no missing data in the 293 questionnaires included in the multivariate linear regression model. There may be some other reasons associated with this "nonsignificant in univariate analysis but significant in multivariate analysis" phenomenon.

Fourth, we used propensity score matching to compare the subset of participants in the completed group with a subset of participants in the uncompleted group who were similar in age and sex [30-32]. The subsets of the completed and uncompleted groups selected using propensity score matching might not be well representative of their respective entire group. As a result, the generalizability of the study results might be of concern.

Fifth, the low adjusted R^2 may be of concern. However, it is expected that the adjusted R^2 value will be low in some fields. For example, fields that attempt to predict human behaviors typically have lower adjusted R^2 values. Given that the nature of this study was to predict human behaviors, it is acceptable that our study results have an adjusted R^2=.1195.

Conclusions

The use of information technology, such as the Internet, to provide health-related information to the general population has grown extremely rapidly in the past decade and will continue to grow at a rapid pace in the future. Our study identified that Internet use is positively associated with frequency of outpatient clinic visits. The contradictory or diverse nature of Internet information might play an important role in the increasing frequency of outpatient clinic visits. In addition, patients' ability to analyze and evaluate health-related information conveyed by a wide variety of media modes for telling whether this information is biased or unbiased may also influence their frequency of making outpatient clinic visits. Future studies may focus on how media literacy affects laypeople's interpretation of health-related information and their health-seeking behaviors.

Acknowledgments

This study was supported by research grants from NTU Triangle (104R104606) and National Taiwan University/National Taiwan University Hospital (104-UN037). The funders did not have any involvement in the data collection, data analysis, data interpretation, writing of this paper, and decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

References

1. Anker AE, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. Patient Educ Couns 2011 Mar;82(3):346-354. [doi: 10.1016/j.pec.2010.12.008] [Medline: 21239134]
2. Taiwan Network Information Center. Taipei, Taiwan, ROC 2012. Press release. URL: http://www.twnic.net.tw/ibnews.php [accessed 2016-12-05] [WebCite Cache ID 6mVSeI2eR]
3. Fox S, Jones S. The social life of health information. Pew Internet & American Life Project. Washington, DC: Pew Research Center; 2009 Jun. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2009/PIP_Health_2009.pdf [accessed 2016-12-05] [WebCite Cache ID 6mWy3IguE]
4. Pew Research Center. Majority of adults look online for health information. Washington, DC: Pew Research Center; 2013 Feb 01. URL: http://www.pewresearch.org/daily-number/majority-of-adults-look-online-for-health-information/ [accessed 2016-12-02] [WebCite Cache ID 6mVSpBdyJ]
5. Bylund CL, Gueguen JA, Sabee CM, Imes RS, Li Y, Sanford AA. Provider-patient dialogue about Internet health information: an exploration of strategies to improve the provider-patient relationship. Patient Educ Couns 2007 Jun;66(3):346-352. [doi: 10.1016/j.pec.2007.01.009] [Medline: 17329062]
6. Bylund CL, Sabee CM, Imes RS, Sanford AA. Exploration of the construct of reliance among patients who talk with their providers about internet information. J Health Commun 2007;12(1):17-28. [doi: 10.1080/10810730601091318] [Medline: 17365346]
7. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
8. Chisholm R, Finnell JT. Emergency department physician internet use during clinical encounters. AMIA Annu Symp Proc 2012;2012:1176-1183 [FREE Full text] [Medline: 23304394]
9. Callahan A, Pernek I, Stiglic G, Leskovec J, Strasberg HR, Shah NH. Analyzing information seeking and drug-safety alert response by health care professionals as new methods for surveillance. J Med Internet Res 2015;17(8):e204 [FREE Full text] [doi: 10.2196/jmir.4427] [Medline: 26293444]
10. Azocar F, McCabe JF, Wetzel JC, Schumacher SJ. Use of a behavioral health web site and service utilization. Psychiatr Serv 2003 Jan;54(1):18. [doi: 10.1176/appi.ps.54.1.18] [Medline: 12509660]
11. Lee C. Does the internet displace health professionals? J Health Commun 2008;13(5):450-464. [doi: 10.1080/10810730802198839] [Medline: 18661387]
12. Ybarra ML, Suman M. Help seeking behavior and the Internet: a national survey. Int J Med Inform 2006 Jan;75(1):29-41. [doi: 10.1016/j.ijmedinf.2005.07.029] [Medline: 16129659]
13. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
14. Takahashi Y, Ohura T, Ishizaki T, Okamoto S, Miki K, Naito M, et al. Internet use for health-related information via personal computers and cell phones in Japan: a cross-sectional population-based survey. J Med Internet Res 2011;13(4):e110 [FREE Full text] [doi: 10.2196/jmir.1796] [Medline: 22169526]
15. Taiwan Social Change Survey. 2016 Nov 11. URL: http://www.ios.sinica.edu.tw/sc/en/home2.php; [accessed 2016-12-02] [WebCite Cache ID 6mVSwvnyr]
16. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
17. Thackeray R, Crookston BT, West JH. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]
18. Naimer S, Mamoot D, Neishtadt A. What you should know about patients who bring a list to their office visit. J Fam Pract 2014 Apr;63(4):E5-E10. [Medline: 24905130]
19. Schwitzer G. How do US journalists cover treatments, tests, products, and procedures? An evaluation of 500 stories. PLoS Med 2008 May 27;5(5):e95 [FREE Full text] [doi: 10.1371/journal.pmed.0050095] [Medline: 18507496]
20. Caulfield T. The commercialisation of medical and scientific reporting. PLoS Med 2004 Dec;1(3):e38 [FREE Full text] [doi: 10.1371/journal.pmed.0010038] [Medline: 15630461]
21. Baron DP. Persistent media bias. J Public Econ 2006 Jan;90(1-2):1-36. [doi: 10.1016/j.jpubeco.2004.10.006]

22. Diem SJ, Lantos JD, Tulsky JA. Cardiopulmonary resuscitation on television. Miracles and misinformation. N Engl J Med 1996 Jun 13;334(24):1578-1582. [doi: 10.1056/NEJM199606133342406] [Medline: 8628340]
23. Chen YY, Chen L, Kao YH, Chu TS, Huang TS, Ko WJ. The over-optimistic portrayal of life-supporting treatments in newspapers and on the Internet: a cross-sectional study using extra-corporeal membrane oxygenation as an example. BMC Med Ethics 2014;15:59 [FREE Full text] [doi: 10.1186/1472-6939-15-59] [Medline: 25081491]
24. Moynihan R, Bero L, Ross-Degnan D, Henry D, Lee K, Watkins J, et al. Coverage by the news media of the benefits and risks of medications. N Engl J Med 2000 Jun 1;342(22):1645-1650. [doi: 10.1056/NEJM200006013422206] [Medline: 10833211]
25. Grilli R, Freemantle N, Minozzi S, Domenighetti G, Finer D. Mass media interventions: effects on health services utilisation. Cochrane Database Syst Rev 2000(2):CD000389. [doi: 10.1002/14651858.CD000389] [Medline: 10796539]
26. Ahmad F, Hudak PL, Bercovitz K, Hollenberg E, Levinson W. Are physicians ready for patients with Internet-based health information? J Med Internet Res 2006;8(3):e22 [FREE Full text] [doi: 10.2196/jmir.8.3.e22] [Medline: 17032638]
27. van Bekkum JE, Hilton S. Primary care nurses' experiences of how the mass media influence frontline healthcare in the UK. BMC Fam Pract 2013;14:178 [FREE Full text] [doi: 10.1186/1471-2296-14-178] [Medline: 24267614]
28. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [Medline: 10396471]
29. Lo SK, Li IT, Tsou TS, See L. Non-significant in univariate but significant in multivariate analysis: a discussion with examples. Chang Gung Medical J 1995;18(2):95-107.
30. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika 1983;70(1):41-55. [doi: 10.1093/biomet/70.1.41]
31. Blackstone EH. Comparing apples and oranges. J Thorac Cardiovasc Surg 2002 Jan;123(1):8-15. [Medline: 11782750]
32. Chen YY, Gordon NH, Connors AF, Garland A, Chang SC, Youngner SJ. Two distinct Do-Not-Resuscitate protocols leaving less to the imagination: an observational study using propensity score matching. BMC Med 2014;12:146 [FREE Full text] [doi: 10.1186/s12916-014-0146-x] [Medline: 25175307]

Abbreviations

TSCS: Taiwan Social Change Survey

Edited by G Eysenbach; submitted 02.01.16; peer-reviewed by WYS Chou, A Callahan, Y Takahashi, R Zheng; comments to author 06.04.16; revised version received 29.06.16; accepted 23.11.16; published 07.12.16.

Please cite as:
Hsieh RW, Chen L, Chen TF, Liang JC, Lin TB, Chen YY, Tsai CC
The Association Between Internet Use and Ambulatory Care-Seeking Behaviors in Taiwan: A Cross-Sectional Study
J Med Internet Res 2016;18(12):e319
URL: http://www.jmir.org/2016/12/e319/
doi:10.2196/jmir.5498
PMID:27927606

Original Paper

Increasing Consumer Engagement by Tailoring a Public Reporting Website on the Quality of Diabetes Care: A Qualitative Study

Maureen A Smith[1,2,3], MPH, MD, PhD; Lauren Bednarz[1,2], MPH; Peter A Nordby[1,2], MA; Jennifer Fink[4,5], PhD; Robert T Greenlee[6], MPH, PhD; Daniel Bolt[7], PhD; Elizabeth M Magnan[8], MD, PhD

[1]Health Innovation Program, School of Medicine and Public Health–Madison, University of Wisconsin, Madison, WI, United States

[2]Department of Population Health Sciences, School of Medicine and Public Health–Madison, University of Wisconsin, Madison, WI, United States

[3]Department of Family Medicine and Community Health, University of Wisconsin, School of Medicine and Public Health–Madison, Madison, WI, United States

[4]Department of Health Informatics, College of Health Sciences, University of Wisconsin–Milwaukee, Milwaukee, WI, United States

[5]Aurora Research Institute, Aurora Health Care, Milwaukee, WI, United States

[6]Center for Clinical Epidemiology, Marshfield Clinic Research Foundation, Marshfield, WI, United States

[7]Department of Educational Psychology, University of Wisconsin–Madison, Madison, WI, United States

[8]Department of Family and Community Medicine, University of California, Davis, Sacramento, CA, United States

Corresponding Author:
Maureen A Smith, MPH, MD, PhD
Health Innovation Program
School of Medicine and Public Health
University of Wisconsin–Madison
800 University Bay Dr
Suite 210
Madison, WI, 53705
United States
Phone: 1 608 262 4802
Fax: 1 888 263 2864
Email: maureensmith@wisc.edu

Abstract

Background: The majority of health care utilization decisions in the United States are made by persons with multiple chronic conditions. Existing public reports of health system quality do not distinguish care for these persons and are often not used by the consumers they aim to reach.

Objective: Our goal was to determine if tailoring quality reports to persons with diabetes mellitus and co-occurring chronic conditions would increase user engagement with a website that publicly reports the quality of diabetes care.

Methods: We adapted an existing consumer-focused public reporting website using adult learning theory to display diabetes quality reports tailored to the user's chronic condition profile. We conducted in-depth cognitive interviews with 20 individuals who either had diabetes and/or cared for someone with diabetes to assess the website. Interviews were audiotaped and transcribed, then analyzed using thematic content analysis.

Results: Three themes emerged that suggested increased engagement from tailoring the site to a user's chronic conditions: ability to interact, relevance, and feeling empowered to act.

Conclusions: We conclude that tailoring can be used to improve public reporting sites for individuals with chronic conditions, ultimately allowing consumers to make more informed health care decisions.

(J Med Internet Res 2016;18(12):e332) doi:10.2196/jmir.6555

KEYWORDS

diabetes; chronic conditions; public reports; patient engagement

Introduction

In spite of evidence that consumers want more information on health care provider performance, there is limited use of current reports to make informed health care decisions [1,2] because only 12% of US adults have consulted online rankings or reviews of clinicians or other physicians [3]. Public reporting on health care performance has been available in the United States for more than two decades and was further enhanced with the Affordable Care Act's [4] creation of a "national strategy for quality improvement through publicly reporting quality performance" [5]. Consequently, health care information transparency initiatives are rapidly increasing. For example, the recently formed Center for Healthcare Transparency has a goal of making information on the relative cost and quality of health care services available to 50% of the US population by 2020 [6]. Yet, studies suggest that consumers do not seek out this information, understand it, trust it, or know how to use it [7,8]. Promising strategies to increase consumer engagement in public reports are to improve the design [9,10], include patient narrative [11-14], and to tailor metrics to reflect the concerns and preferences of individual consumers [15]. It has been shown that the absence of tailoring in public reports makes them unlikely to succeed [16,17] and that consumers are deterred by the content and design of current reports that lack tailoring to their individual needs [18]. Tailoring, or personalizing health information, means creating communicative information about a given individual to determine what specific content they will receive, the contexts or frames surrounding the content, by whom it will be presented, and the channels through which it will be delivered [19].

Consumers with multiple chronic conditions are a priority population for tailoring public reports because they have a continuing need to know how to best manage their health conditions to avoid complications and improve their health [20]. It is increasingly recognized that existing reporting initiatives do not support decision making for these consumers even though persons with chronic conditions may be the population most interested in public reports [7,21]. Those with diabetes mellitus are particularly representative of this group because more than 90% have multiple chronic conditions (diabetes plus at least one more condition) [22]. Additionally, persons with diabetes may be more receptive to publicly reported information on quality [23] because of their emotional connection to their disease, awareness of symptoms and consequences, and information-seeking behaviors.

We examined whether tailoring existing public reports on health care quality to persons with diabetes and co-occurring chronic conditions would increase their engagement with the reports. Using four principles of adult learning [24], we adapted a consumer-focused website that publicly reports the quality of diabetes care using a hypothetical patient narrative [25] by tailoring the information presented more specifically to each person's co-occurring chronic conditions. We shared the original and revised websites with consumers and conducted semistructured interviews to obtain feedback on the changes. We expected that greater customization of content tailored to the individual's conditions would increase engagement with this public reporting website [26].

Methods

Study Design

As recommended in the 2025 vision for public reporting [20], we engaged consumers in developing and testing public reports as a mechanism to determine what is most useful and meaningful to them. We adapted an existing public reporting website that uses a novel storytelling format to explore the health issues and health care options of "Helen," who has diabetes [27]. In its existing format, consumers can follow and learn from Helen's story, compare the performance of more than 20 health systems on metrics representing the quality of diabetes care within each health system, and find useful tips about being a better health care consumer. For our study, we redesigned the website to tell the story of "Karen," who has diabetes and other health issues (Multimedia Appendix 1 Figure A1). The website redesign was interactive and allowed the content to be tailored to each user to reflect health system performance metrics for individuals with similar chronic conditions. Next, to assess the websites, we conducted in-depth cognitive interviews with 20 individuals who either had diabetes and/or cared for someone with diabetes [28]. During the interviews, we asked participants to navigate through the existing website with Helen's story and the redesigned website with Karen's story. Interviews were audiotaped and transcribed, then analyzed using thematic content analysis.

Developing a Website That Tailors Information to the User

To create a website that tailors information for the user, we used four principles of adult learning: (1) adult learners' experience, (2) involved adult learner, (3) problem-centered, and (4) relevance and impact to learners' lives [24]. Our adaptations to address these principles included (1) creating a new story about Karen who has other health issues in addition to diabetes, (2) allowing the user to answer four questions about their own chronic conditions, (3) using the four questions to create a chronic condition profile for the user (Multimedia Appendix 1 Figure A2), and (4) generating a report comparing quality among health systems for that user's chronic condition profile.

Sampling

To recruit, flyers advertising the study were posted at local libraries, grocery stores, community centers, senior centers, university buildings, and a local hospital. Additionally, we placed an ad in a local paper. Only individuals who had diabetes and/or cared for an individual with diabetes were eligible. Those interested in participating in the study were screened via telephone by the University of Wisconsin Survey Center. We expected our study population to be older, white, and college educated based on the demographics of the local population. Once we reached 20 participants, we closed enrollment into the study. We based our sample size on the goal of saturation (ie, little new is being learned) [29,30]. Each participant received a US $50 stipend on the completion of their interview.

Data Collection

Between April and May 2015, we conducted 20 one-on-one, hour-long, in-depth cognitive interviews. During the interviews, the participants navigated through the existing consumer-focused Web pages with Helen's story while being asked scripted, open-ended questions related to (1) navigation of content, (2) validity of the content, and (3) understanding of the displays. The participants were then instructed to navigate through the newly adapted Web pages with Karen's story while being asked the same scripted, open-ended questions plus questions about (4) their understanding of information presented for the chosen chronic condition profile, and (5) perceived value of the health care quality report presented for the user's chronic condition profile. The interviews were audio recorded and transcribed. Short summary thoughts and notes were drafted after each interview.

Data Analysis

We used thematic content analysis to identify, analyze, and report themes within our data. As a guide, we applied Braun and Clarke's "six phases of analysis" in which coders (1) familiarize themselves with the data, (2) generate codes, (3) search for themes among codes, (4) review themes, (5) define and name themes, and (6) produce the report [31]. Data analysis was performed by two research assistants who independently coded each transcript in full and met regularly to discuss coding and reach interrater reliability. Agreed-on codes were then examined to identify reoccurring themes in our dataset.

Results

Our study population (n=20) was primarily female (n=13) with racial minorities making up 30% (6/20) of our sample (Table 1). The median age range of our participants was between 41 and 50 years old. The majority (60%, 12/20) of participants had health insurance either through Medicare (n=4), Medicaid (n=8), or a mix of Medicaid and/or Medicare supplemented with private insurance (n=5). Participants included 8 persons with diabetes, 8 caregivers of persons with diabetes, and 4 participants who were both persons with diabetes and caregivers for persons with diabetes.

Themes

The findings from our thematic content analysis support the concept that tailoring a public reporting website to a user's chronic condition profile can increase consumer engagement. Participants found that the ability to interact (Theme 1) with the website to generate a chronic condition profile personalized their experience, whereas the addition of Karen's story to the website increased the relevance (Theme 2) of the site because it presented the voice of an individual who has "other health issues" in addition to diabetes (Textbox 1). The presence of these two themes—ability to interact and relevance—led to participants' feeling empowered to act (Theme 3) to improve their health care and health.

Textbox 1. Themes identified and relevant quotes from interviews with consumers' while navigating a public reporting website on health care quality tailored to a user's chronic condition profile compared to a nontailored website.

Theme 1: Ability to interact

- "It's nice that, like, you can kind of put your conditions in and see how people are performing."
- "Well, right off the bat this is more personal because it's involving my opinion or self-data. So, it's more personal."
- "They should maybe have a box where you could actually type in comments because you might want to specify what issues you're dealing with. I would like to provide more information."

Theme 2: Relevance

- "...it's really good to hear personal stories of people going through the same things—just to be able to relate to that..."
- "She [Karen] said that she had other health problems in addition to diabetes, which a lot of people have and can probably relate to."
- "This one's [Karen's] a lot better. I just realized that this one [Helen's] tells you nothing, nothing at all. This one right here is pretty much breaking it down for me. Karen's—I can relate to her more."

Theme 3: Feeling empowered to act

- "I think that it's just a way of knowing which health system would take care of you the best to manage your disease."
- "Well, if I'm changing insurance, and I have the ability to choose a doctor, I want to choose a doctor that knows what they're doing."
- "It's important to do your research, because that is true, some doctors, hospitals, and clinics are better when they're dealing with patients with multiple issues...in addition to diabetes. I didn't know that."

Table 1. Participant characteristics (n=20).

Participant characteristics	n (%)
Age range (years)	
18-30	5 (25)
31-40	2 (10)
41-50	6 (30)
51-60	6 (30)
61-70	1 (5)
Sex	
Female	13 (65)
Male	7 (35)
Race	
White	13 (65)
Black or African American	5 (25)
Other	1 (5)
Unidentified	1 (5)
Income (US $)	
<$25,000	8 (40)
$25,000-$40,000	4 (20)
$40,000-$65,000	6 (30)
$75,000-$100,000	1 (5)
Unknown	1 (5)
Health insurance status	
Private	8 (40)
Medicare	3 (15)
Medicaid	4 (20)
Mix of Medicaid and/or Medicare with Private	5 (25)
Employment status	
Employed	13 (65)
Unemployed	7 (35)
Highest level of education	
High school diploma or equivalent	3 (15)
Some college	7 (35)
Associate’s degree	2 (10)
Bachelor’s degree	7 (35)
Master’s degree	1 (5)

Ability to Interact

The process of answering questions about their chronic conditions allowed participants to interact directly with the site to create their chronic condition profile and, based on this profile, display information on which health systems perform best for a “patient like them.”

Participants saw value in interacting with the site to create a profile based on their conditions. One participant stated that it was nice to be able to select your conditions and receive more “personal” data. In fact, some participants would have liked the ability to interact with the site on an even more personal and comprehensive level via a comment box where they could be more specific about their health and conditions. Another participant suggested having a checklist of all possible conditions for the user to self-select. Our findings suggest that the more a user is allowed to interact on a website that reports on health care quality, the more engaged they may become with website and ultimately the information that the website provides.

Overwhelmingly, participants in the study found the website easy to navigate (90%, 18/20). Because the primary goal of the website was for participants to be able to create a tailored report on health care quality by self-selecting their conditions to create a chronic condition profile, the ease of navigating the website was an integral piece of allowing that interaction to take place.

Relevance

Karen's story (multiple chronic conditions in addition to diabetes) was viewed by the participants as being more relevant because Karen faced challenges similar to their own. Adding Karen's story to the website provided a relatable narrator for the participants in our study. It was easy for them to "see themselves" in Karen's story because she, like them, lives with diabetes and other chronic conditions. As one participant remarked, "There's more to the issue of complicated diabetes...I don't know many diabetics that don't have other health issues."

Participants viewed receiving a score on how health systems perform for an individual with diabetes and other health issues similar to theirs, as opposed to a score for all patients with diabetes, as significant to them as an individual. Providing a platform for consumers to identify their conditions, to create a profile, and receive a score on health system performance that is personalized for them allows consumers to "see" themselves in the data. This increased the relevance of the health care quality report being communicated because it was specifically related to the consumer and their chronic condition profile. When asked if the website tailored to their chronic condition profile provided more value compared to the untailored website, 13 of 20 participants said yes, 3 participants did not comment, and 4 found no additional value.

Feeling Empowered to Act

The process of self-selecting their conditions and receiving personalized data on how health systems perform for a "patient like them" created a sense of empowerment for the participants. This feeling of empowerment was stimulated by providing a report that showed which health systems perform best in managing patients with their similar chronic condition profile.

Participants suggested that knowing how health systems perform for individuals similar to themselves may influence where they would seek out care. For example, one participant stated they would be more likely to see a provider if the data showed they were good at treating patients with a similar profile to theirs. Participants also thought they would find this information particularly valuable if they were changing insurance. This highlights the importance of making public reports available to consumers at times when they are most likely to make a decision about their health care including health care insurance coverage (eg, open enrollment periods).

Lastly, Karen's narrative served as a prescriptive guide for participants on how they should interact with their provider to better manage their health. One participant stated, "If you don't understand something, you have to be honest. And you have to say (to your provider), you know, 'I don't understand.'" Other participants felt empowered by the narrative to initiate conversations with their provider and become more of an active participant in their health.

Discussion

The development of personalized or customizable report cards has been identified as an important priority to the success of consumer-focused quality reporting [32], although there has been a dearth of evidence on the consumer perspective on personalization [18]. By engaging consumers, we found that tailoring a website that publicly reports on health care quality to a consumer's chronic condition profile increased consumer engagement with the website in specific ways. The themes that emerged from our analysis suggest that consumers value a public reporting website that they can interact with, that they feel is relevant to their situation, and that provides them with a feeling of empowerment to support action. Our findings support that tailoring is indeed a potential strategy for increasing consumer engagement in public reports as suggested by Huckman [15].

Our specific approach of tailoring a public reporting website based on the principles of adult learning as a framework represents a useful guide to increase consumer engagement. Multiple different strategies have been proposed to increase consumer engagement in public reporting, such as providing cost data, having patient comments and stories on the sites, better design, and providing information that is more relevant to the user [18,33]. However, there has been a lack of an overarching framework to organize and prioritize these multiple strategies. Our framework is consistent with consumer requests to have public reporting websites provide more personally relevant information on health care quality for patients with similar health conditions [34]. Also supporting our framework, we found that the participants in our study wish to engage further with the website we adapted. For example, participants expressed a desire to enter more detailed information about their health conditions in order to create an even more personalized quality report. This suggests that ensuring the delivery of public reports on health care quality are adequately personalized to the consumer is important to consider because personalization is an important component of effectiveness [35,36]. In other settings, tailoring educational materials to individual consumers increases the chances that the information will be read and remembered, saved and discussed with others, perceived as interesting, and personally relevant [37].

There are multiple ways to tailor health information, such as matching to race, ethnicity, gender, or age [34], but we chose to tailor reports for persons with multiple chronic conditions for a variety of reasons. Currently in the United States, the majority of health care utilization decisions are made by the quarter of individuals who have multiple chronic conditions, comprising approximately 66% of health care expenditures in the United States [38]. Having chronic conditions is a powerful motivator to seek out the best health care available because individuals with chronic conditions have a continuing need to know how to best manage their health conditions to avoid complications or deterioration, minimize symptoms, and improve their health [21]. In addition, individuals with chronic conditions see multiple providers of potentially differing quality in the face of complex medical needs [21]. By allowing consumers to tailor the reports to their chronic condition profile, we are delivering content that is relevant to their interests and needs [26].

Our study has several limitations. First, although building on the design of a preexisting consumer website as a basis for our assessment reduced development time and facilitated real-world testing, it constrained our options in presentation of data. Second, we had an overrepresentation of low-income and racial minorities in our sample compared to the expected population. However, our sample represents a high-priority population as a proposed strategy to reduce health disparities and improve health outcomes by making health information, such as public reports available to racial and ethnic minorities [39].

Although consumers want more transparency and information on health care provider performance, they do not use existing public websites that provide this information. We found that redesigning a public reporting website using adult learning theory is one way to potentially increase consumer engagement with these sites. Participants in our study valued their ability to interact with the website, felt the information on the site had relevance to them, and felt empowered to act based on the information provided. Our approach provides specific guidance on changing the content and format of public reports to engage and inform consumer decisions. A potential area for future study is how to tailor health system performance metrics to additional consumer attributes that might affect the quality of care (eg, race/ethnicity), to the relevance of diabetes in a person's perception of health, or to awareness of other comorbidities, and how these might affect diabetes care.

Acknowledgments

Funding/Financial support: This project was funded by grant R21 HS021899 from the Agency for Healthcare Research and Quality. Additional support was provided by the Health Innovation Program, the UW School of Medicine and Public Health from The Wisconsin Partnership Program, and the Community - Academic Partnerships core of the University of Wisconsin Institute for Clinical and Translational Research (UW ICTR) through the National Center for Advancing Translational Sciences (NCATS), grant UL1TR000427. Dr Magnan was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through grant #UL1 TR001860 and linked award KL2 TR001859.

Role of the funder: The study sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

We would like to thank the Wisconsin Collaborative for Healthcare Quality and the University of Wisconsin Survey Center for their assistance in the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Public reporting website storyboards.

[PDF File (Adobe PDF File), 414KB - jmir_v18i12e332_app1.pdf]

References

1. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. JAMA 2000 Apr 12;283(14):1866-1874. [Medline: 10770149]
2. Lagu T, Lindenauer PK. Putting the public back in public reporting of health care quality. JAMA 2010 Oct 20;304(15):1711-1712. [doi: 10.1001/jama.2010.1499] [Medline: 20959582]
3. Fox S. The social life of health information. 2011. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2016-08-23] [WebCite Cache ID 6jz2IbgTJ]
4. Eichner J, Blumenthal D, editors. Medicare in the 21st Century: Building a Better Chronic Care System. Washington, DC: National Academy of Social Insurance; 2003 Jan. URL: https://www.nasi.org/sites/default/files/research/Chronic_Care_Report.pdf [WebCite Cache ID 6jz3KI1pQ]
5. Health Affairs. Health policy brief: public reporting on quality and costs. 2012 Mar 08. URL: http://www.healthaffairs.org/healthpolicybriefs/brief.php?brief_id=65 [accessed 2016-08-23] [WebCite Cache ID 6jz4LBUVo]
6. Center for Healthcare Transparency. 2014. URL: http://www.nrhi.org/uploads/center-for-healthcare-transparency-vision-and-principles.pdf [accessed 2016-08-23] [WebCite Cache ID 6jz3SpNaw]
7. Marshall MN, Shekelle PG, Davies HT, Smith PC. Public reporting on quality in the United States and the United Kingdom. Health Aff (Millwood) 2003;22(3):134-148 [FREE Full text] [Medline: 12757278]
8. Robinowitz DL, Dudley RA. Public reporting of provider performance: can its impact be made greater? Annu Rev Public Health 2006;27:517-536. [doi: 10.1146/annurev.publhealth.27.021405.102210] [Medline: 16533128]

9. Damman OC, van den Hengel YK, van Loon AJ, Rademakers J. An international comparison of web-based reporting about health care quality: content analysis. J Med Internet Res 2010;12(2):e8 [FREE Full text] [doi: 10.2196/jmir.1191] [Medline: 20439252]
10. Sander U, Emmert M, Dickel J, Meszmer N, Kolb B. Information presentation features and comprehensibility of hospital report cards: design analysis and online survey among users. J Med Internet Res 2015;17(3):e68 [FREE Full text] [doi: 10.2196/jmir.3414] [Medline: 25782186]
11. Greaves F, Millett C. Consistently increasing numbers of online ratings of healthcare in England. J Med Internet Res 2012;14(3):e94 [FREE Full text] [doi: 10.2196/jmir.2157] [Medline: 22742977]
12. 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]
13. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. J Med Internet Res 2013;15(2):e24 [FREE Full text] [doi: 10.2196/jmir.2360] [Medline: 23372115]
14. Bardach NS, Hibbard JH, Greaves F, Dudley RA. Sources of traffic and visitors' preferences regarding online public reports of quality: web analytics and online survey results. J Med Internet Res 2015;17(5):e102 [FREE Full text] [doi: 10.2196/jmir.3637] [Medline: 25934100]
15. Huckman RS, Kelley MA. Public reporting, consumerism, and patient empowerment. N Engl J Med 2013 Nov 14;369(20):1875-1877. [doi: 10.1056/NEJMp1310419] [Medline: 24131139]
16. Joseph Christine L M, Peterson E, Havstad S, Johnson CC, Hoerauf S, Stringer S, Asthma in Adolescents Research Team. A web-based, tailored asthma management program for urban African-American high school students. Am J Respir Crit Care Med 2007 May 1;175(9):888-895 [FREE Full text] [doi: 10.1164/rccm.200608-1244OC] [Medline: 17290041]
17. Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. JAMA 1998 May 27;279(20):1638-1642. [Medline: 9613914]
18. Hussey PS, Luft HS, McNamara P. Public reporting of provider performance at a crossroads in the United States: summary of current barriers and recommendations on how to move forward. Med Care Res Rev 2014 Oct;71(5 Suppl):5S-16S. [doi: 10.1177/1077558714535980] [Medline: 24871273]
19. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
20. Shaller D, Kanouse DE, Schlesinger M. Context-based strategies for engaging consumers with public reports about health care providers. Med Care Res Rev 2014 Oct;71(5 Suppl):17S-37S. [doi: 10.1177/1077558713493118] [Medline: 23819945]
21. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. J Med Internet Res 2013;15(8):e187 [FREE Full text] [doi: 10.2196/jmir.2702] [Medline: 23985220]
22. Magnan EM, Palta M, Mahoney JE, Pandhi N, Bolt DM, Fink J, et al. The relationship of individual comorbid chronic conditions to diabetes care quality. BMJ Open Diabetes Res Care 2015;3(1):e000080 [FREE Full text] [doi: 10.1136/bmjdrc-2015-000080] [Medline: 26217492]
23. Robert Wood Johnson Foundation. Consumer Decision Points of Health Care Information. Princeton, NJ: Robert Wood Johnson Foundation; 2010 Jul. URL: http://www.rwjf.org/content/dam/farm/reports/reports/2010/rwjf69380 [accessed 2016-12-12] [WebCite Cache ID 6mhehUSJS]
24. Knowles M. In: Wilton C, editor. Knowles MS. The modern practice of adult education: from pedagogy to andragogy (Rev. and Updated. ed.) [Wilton, Conn.]. Chicago: Association Press; Follett Pub. Co; 1980.
25. WCHQ Ambulatory Measure Specification: Diabetes Care Performance Measures. Middleton, WI: Wisconsin Collaborative for Healthcare Quality; 2011.
26. Damberg CL, McNamara P. Postscript: research agenda to guide the next generation of public reports for consumers. Med Care Res Rev 2014 Oct;71(5 Suppl):97S-107S. [doi: 10.1177/1077558714535982] [Medline: 24860123]
27. My Health Wisconsin. Getting good care. URL: http://www.myhealthwi.org/Resources/GettingGoodCare.aspx [accessed 2016-12-06] [WebCite Cache ID 6mYiqRaiF]
28. Bradley CJ, Given CW, Dahman B, Fitzgerald TL. Adjuvant chemotherapy after resection in elderly Medicare and Medicaid patients with colon cancer. Arch Intern Med 2008 Mar 10;168(5):521-529. [doi: 10.1001/archinternmed.2007.82] [Medline: 18332299]
29. Hibbard J, Sofaer S. Best Practices in Public Reporting No. 1: How To Effectively Present Health Care Performance Data To Consumers. Rockville, MD: Agency for Healthcare Research and Quality; 2010. URL: http://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/pubrptguide1/pubrptguide1.pdf [accessed 2016-08-23] [WebCite Cache ID 6jz3bVcNJ]
30. Sofaer S, Hibbard J. Best Practices in Public Reporting No. 2: Maximizing Consumer Understanding of Public Comparative Quality Reports: Effective Use of Explanatory Information. Rockville, MD: Agency for Healthcare Research and Quality; 2010. URL: http://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/pubrptguide2/pubrptguide2.pdf [accessed 2016-08-23] [WebCite Cache ID 6jz3eMl92]

31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. Sinaiko AD, Eastman D, Rosenthal MB. How report cards on physicians, physician groups, and hospitals can have greater impact on consumer choices. Health Aff (Millwood) 2012 Mar;31(3):602-611 [FREE Full text] [doi: 10.1377/hlthaff.2011.1197] [Medline: 22392672]
33. Bardach N, Hibbard J, Dudley R. Users of Public Reports of Hospital Quality: Who, What, Why and How? An Aggregate Analysis of 16 Online Public Reporting Web Sites and Users' and Experts' Suggestions for Improvement. Rockville, MD: Agency for Healthcare Research and Quality; 2011. URL: http://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/value/pubreportusers/pubreportusers.pdf [accessed 2016-08-23] [WebCite Cache ID 6jz3ggZPL]
34. Marshall MN, Romano PS, Davies HT. How do we maximize the impact of the public reporting of quality of care? Int J Qual Health Care 2004 Apr;16 Suppl 1:i57-i63 [FREE Full text] [doi: 10.1093/intqhc/mzh013] [Medline: 15059988]
35. Weymann N, Dirmaier J, von WA, Kriston L, Härter M. Effectiveness of a Web-based tailored interactive health communication application for patients with type 2 diabetes or chronic low back pain: randomized controlled trial. J Med Internet Res 2015;17(3):e53 [FREE Full text] [doi: 10.2196/jmir.3904] [Medline: 25736340]
36. Kreuter M, Farrell D, Olevitch L, Brennan L. Tailoring Health Messages: Customizing Communication with Computer Technology. Mahwah, NJ: Lawrence Erlbaum Associates; 2000:270.
37. Kreuter MW, Bull FC, Clark EM, Oswald DL. Understanding how people process health information: a comparison of tailored and nontailored weight-loss materials. Health Psychol 1999 Sep;18(5):487-494. [Medline: 10519465]
38. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007 Dec;22 Suppl 3:391-395 [FREE Full text] [doi: 10.1007/s11606-007-0322-1] [Medline: 18026807]
39. Yonek JC, Mittler JN, Hasnain-Wynia R. Why and how six aligning forces for quality communities have focused on reducing disparities. Med Care Res Rev 2014 Oct;71(5):435-449. [doi: 10.1177/1077558714533826] [Medline: 24821676]

Edited by G Eysenbach; submitted 26.08.16; peer-reviewed by U Sander, M Schlesinger; comments to author 02.11.16; revised version received 14.11.16; accepted 23.11.16; published 21.12.16.

Please cite as:
Smith MA, Bednarz L, Nordby PA, Fink J, Greenlee RT, Bolt D, Magnan EM
Increasing Consumer Engagement by Tailoring a Public Reporting Website on the Quality of Diabetes Care: A Qualitative Study
J Med Internet Res 2016;18(12):e332
URL: http://www.jmir.org/2016/12/e332/
doi:10.2196/jmir.6555
PMID:28003173

Original Paper

Treatment of Acute Coronary Syndrome by Telemedically Supported Paramedics Compared With Physician-Based Treatment: A Prospective, Interventional, Multicenter Trial

Jörg C Brokmann[1], MD; Clemens Conrad[2], MD; Rolf Rossaint[2], MD; Sebastian Bergrath[2], MD; Stefan K Beckers[2], MD; Miriam Tamm[3], Dipl-Math; Michael Czaplik[2], MD; Frederik Hirsch[2], MD

[1]Emergency Department, Rheinisch-Westfälische Technische Hochschule, University Hospital RWTH Aachen, Aachen, Germany
[2]Department of Anaesthesiology, Rheinisch-Westfälische Technische Hochschule, University Hospital RWTH Aachen, Aachen, Germany
[3]Department of Medical Statistics, Rheinisch-Westfälische Technische Hochschule, University Hospital RWTH Aachen, Aachen, Germany

Corresponding Author:
Frederik Hirsch, MD
Department of Anaesthesiology
Rheinisch-Westfälische Technische Hochschule
University Hospital RWTH Aachen
Pauwelsstrasse 30
Aachen, 52074
Germany
Phone: 49 0241 88179
Fax: 49 024182406
Email: fhirsch@ukaachen.de

Abstract

Background: Prehospital treatment of acute coronary syndrome (ACS) in German emergency medical services (EMSs) is reserved for EMS physicians due to legal issues.

Objective: The objective of this prospective, interventional, multicenter trial was to evaluate the quality of telemedically-delegated therapy and the possible complications in patients with ACS.

Methods: After approval by the ethics committee and trial registration, a one-year study phase was started in August 2012 with 5 ambulances, telemedically equipped and staffed with paramedics, in 4 German EMS districts. The paramedics could contact an EMS-physician–staffed telemedicine center. After initiation of an audio connection, real-time data transmission was automatically established. If required, 12-lead electrocardiogram (ECG) and still pictures could be sent. Video was streamed from inside each ambulance. All drugs, including opioids, were delegated to the paramedics based on standardized, predefined algorithms. To compare telemedically-delegated medication and treatment in ACS cases with regular EMS missions, a matched pair analysis with historical controls was performed.

Results: Teleconsultation was performed on 150 patients having a cardiovascular emergency. In 39 cases, teleconsultation was started due to suspected ACS. No case had a medical complication. Correct handling of 12-lead ECG was performed equally between the groups (study group, n=38 vs control group, n=39, $P>.99$). There were no differences in correct handling of intravenous administration of acetylsalicylic acid, heparin, or morphine between both the groups (study group vs control group): acetylsalicylic acid, n=31 vs n=33, $P=.73$; unfractionated heparin, n=34 vs n=33, $P>.99$; morphine, n=29 vs n=27, $P=.50$. The correct handling of oxygen administration was significantly higher in the study group (n=29 vs n=18, $P=.007$).

Conclusions: Telemedical delegation of guideline conform medication and therapy by paramedics in patients with ACS and was found to be feasible and safe. The quality of guideline-adherent therapy was not significantly different in both the groups except for the correct administration of oxygen, which was significantly higher in the study group.

Trial Registration: Clinicaltrials.gov NCT01644006; http://clinicaltrials.gov/ct2/show/NCT01644006 (Archived by WebCite at http://www.webcitation.org/6mPam3eDy).

(J Med Internet Res 2016;18(12):e314) doi:10.2196/jmir.6358

KEYWORDS

acute coronary syndrome; prehospital emergency care; telemedicine; telehealth; myocardial infarction

Introduction

The in-hospital mortality rate of ST Segment Elevation Myocardial Infarction (STEMI) could be significantly reduced by the modern reperfusion therapy and improved secondary prophylaxis. However, the overall mortality remains unchanged because two-thirds of deaths occur during the prehospital phase—usually due to lethal arrhythmias that are triggered by ischemic events [1]. Therefore, an appropriate therapy must be initiated in the early prehospital phase, and the time of admission must be as small as possible to improve survival. This is also important in non–STEMI-acute coronary syndrome (NSTE-ACS) as the incidence of NSTE-ACS rises further [2,3].

Telemedical support for diagnosis and therapy in acute coronary syndromes (ACSs) has been established for quite some time in emergency medical services (EMS). Several telemetry projects have demonstrated feasibility [4-7] in the transmission of a 12-lead electrogram (ECG) or its images to a cardiologist. The cardiologist receiver can then consecutively assess the potential to improve survival and outcome in STEMI [7]. It even has been shown that an artificial neural network could predict STEMI and the need of acute percutaneous coronary intervention (PCI) in ambulance ECGs [8]. However, all of these studies assessed the impact of 12-lead ECG transmission on in-hospital parameters and outcome in STEMI patients. Evidence regarding telemedical concepts in non-STEMI ACS patients is mostly lacking.

Germany uses a two-tiered EMS system with paramedic-staffed ambulances and additional EMS with physician response units. Like some other European countries, only an EMS physician can administer the medication required in ACS. Except for nitrates, paramedics cannot dispense drugs. In cases where primary dispatch was merely a paramedic-staffed ambulance and the EMS physician was notified later, the time to arrival and thus the administration of medication may be unacceptably long. This effect is exacerbated in more rural areas and by an occasional lack of physician-staffed EMS [9,10].

Against this background, a mobile telemedicine system was developed during a first research project from 2007 to 2010. Two comprehensive simulation studies demonstrated that real-time telemedical support by experienced remote EMS physicians leads to improved quality of care in STEMI. Even telemedically supported paramedics were able to handle emergency care on their own with a comparable performance compared with on-scene physicians [11,12]. The main findings during the project, preceding to the one described here, were the feasibility of prehospital teleconsultation in general and the improvement of data transmission in acute stroke [13,14]. This succeeding research project allowed further technical and organizational development. The telemedical equipment was made more practicable and paramedic-staffed ambulances were equipped with this system. Thus, the concept allowed medications to be given by paramedics in ACS when supervised by an experienced EMS physician—the so-called tele-EMS physician. In this prospective, interventional trial, we assessed the quality of prehospital emergency care in patients suffering from ACS (measured by adherence to international guidelines) when using a telemedically supported EMS. Using matched pairs, we compared our findings with a comparable historical control group treated by our regular physician-based EMS.

Methods

This prospective, interventional, multicenter trial was conducted from August 1, 2012 to July 31, 2013 within the research project “TemRas” (telemedical rescue assistance system).

Study Setting

In 4 different EMS districts in Germany, 5 paramedic-staffed ambulances were stepwise equipped with a multifunctional telemedicine system. Three of the 4 EMS districts were rural and one was urban (Table 1). As described elsewhere, all participating EMS physicians in the function of a “tele-EMS physician” were trained prior to the start of the intervention phase [15]. All participating tele-EMS physicians had a minimum experience of 3 years in anesthesia and critical care as well as broad experience as an EMS physician. The paramedics involved (N=178) ran through a standardized eight-hour training program before the project started. The goal was to learn the use of the technical system, the medical concept of teleconsultation including indications for teleconsultation (eg, ACS), and communication skills [15].

Table 1. Demographics and structure of the participating emergency medical service (EMS) districts.

Demographics	Aachen (urban)	Heinsberg (rural)	Dueren (rural)	Euskirchen (rural)
Population	248,137	256,546	267,712	190,591
Area (km^2)	160.8	628.0	941.4	1248.7
Ambulances				
24-h Ambulances (telemedically equipped[a])	6 (2)	7 (1)	11 (1)	9 (1)
Daytime ambulances	2	3	2	2
EMS[b] physician units	2+1[c]	4	4+1[d]	3
Ambulance emergency missions/year	22,984	14,346	20,302	15,108
EMS physician missions/year	7898	7786	9057	5317
Hospitals	4	4	5	3
Stroke units	1	1	1	1
Level 1 trauma center	1	-	-	-
24-h Cardiac catheterization laboratory	1	1	1	1
Number of cases (n)	18	6	3	12

[a] telemedically equipped for 24 hours, operation of teleconsultation center: 7:30 am-4:30 pm.

[b] EMS: emergency medical service.

[c] additional EMS physician at home can be picked up by the fire department.

[d] daytime EMS physician unit: 7 am-4 pm.

Trial Registration and Ethical Issues

The trial was registered before the intervention phase (clinicaltrials.gov NCT01644006). The study was approved by the local ethics committee (University Hospital Rheinisch-Westfälische Technische Hochschule Aachen, Germany, registration number EK 191/11). The written informed consent was waived by the ethics committee due to the emergency setting; however, all alert patients had to give verbal approval to data and video transmission before teleconsultation. Data from the preintervention phase was analyzed retrospectively via prospectively planned outcomes of the intervention phase. Multimedia Appendix 1 shows the trial protocol.

Technical System

An MRx monitor-defibrillator unit (Philips Healthcare, Andover, MA, USA) connected to a portable data transmission unit (peeqBox, P3 communications, Aachen, Germany) was used by the paramedics to establish parallelized and encrypted data and audio connection via different mobile networks. The audio channel was accomplished with 2 headsets (Voyager Pro HD, Plantronics, Santa Cruz, CA, USA) connected via Bluetooth to the mobile transmission unit. Real-time vital data transmission (numerical values and curves, Einthoven leads I-III continuously) occurred automatically after initiation of a teleconsultation. If required, 12-lead ECGs could be transmitted. Still pictures could be sent using a Bluetooth connection between a mobile phone (HTC Sensation XE, High Tech Computer Corporation, Taoyuan, Taiwan) and the data transmission unit. The tele-EMS physician could start a video streaming from a ceiling camera of the ambulance when the patient was inside in the ambulance. On the telemedical workstation, real-time vital data was displayed using the IntelliVue Information Centre (Philips Healthcare, Boeblingen, Germany), and 12-lead-ECGs were displayed on the HeartStart Telemedicine Viewer (Philips Healthcare, Andover, MA, USA). All other software components in the telemedicine center were specifically developed as part of the research project. In cases of ACS, an established algorithm could be displayed on the workstation to support the tele-EMS physician in guideline-adherent therapy recommendations including the steps needed to diagnose and dose medication as recommended by national and international guidelines. A detailed technical description of the telemedicine system was already described elsewhere [16].

Interventions

The decision to initiate a teleconsultation was solely made by the on-scene paramedics. All participating EMS districts ran their own dispatch centers. Local protocols for ambulance and EMS physician units' alarm were not changed before or within the project period. The telemedicine system was used in addition to the regular EMS system. If the EMS dispatch centers suspected ACS, then an EMS physician response unit was sent in accordance to the special provisions of law for paramedic-staffed ambulance. Therefore, only cases where the initial notification was different or cases that bridged the arrival of an EMS physician unit could be managed with telemedical care. The additional telemedicine service was available during workdays from 7:30 am to 4:30 pm during the first 4 months of the intervention phase and from 7:30 am to 8:00 pm for the rest of the trial phase due to restricted funding. No teleconsultation system existed during the preintervention period.

Outcomes and Data Sources

The following data sources were used: the teleconsultation protocols, paper-based EMS protocols of the participating EMS stations, and data of the local EMS dispatch centers. The quality of prehospital care of ACS based on national and international published guidelines [17-21] for non–STEMI-ACS and STEMI was analyzed as the primary outcome measure. Guideline adherence was measured by correct handling of the following measures: 12-lead ECG completed; administration of acetylsalicylic acid≥250mg iv, independent of long-term therapy of the patient; administration of heparin 70 IU/kg iv, maximum 5000 IU, but waiver in the presence of effective anticoagulation in patient's long-term therapy; administration of morphine if numerical rating scale (NRS)≥3 (repeatedly 3-5 mg iv until NRS≤3); and administration of oxygen if oxygen saturation≤95%. Treatment was not carried out in cases of any allergies or contraindications. If so, this was measured as correct handling.

A standardized report for any potential problems or adverse events was completed daily by the tele-EMS physician in charge. The adverse events were defined as follows: respiratory or circulatory insufficiency due to administered medications with a need for intervention or allergic reaction due to administered medications. Multimedia Appendix 2 shows the original dataset for further statistical evaluation.

Sample Size

The trial had a restricted funding period of one year. Because this was the first study of prehospital telemedically-delegated diagnosis and therapy including medication in patients with ACS in this setting, a power analysis and formal sample size calculation could not be performed before this pilot study. A possible sample size was estimated by an expert group and documented in the trial registration. The anticipated enrolment of the sample size was N=180.

Inclusion and Exclusion Criteria

Patients were included if ACS was diagnosed, verbal consent for teleconsultation was obtained, and were aged ≥18 years. Patients who refused consent to teleconsultation or had no suspected ACS were excluded. Figure 1 shows the study flowchart.

Matched Pairs

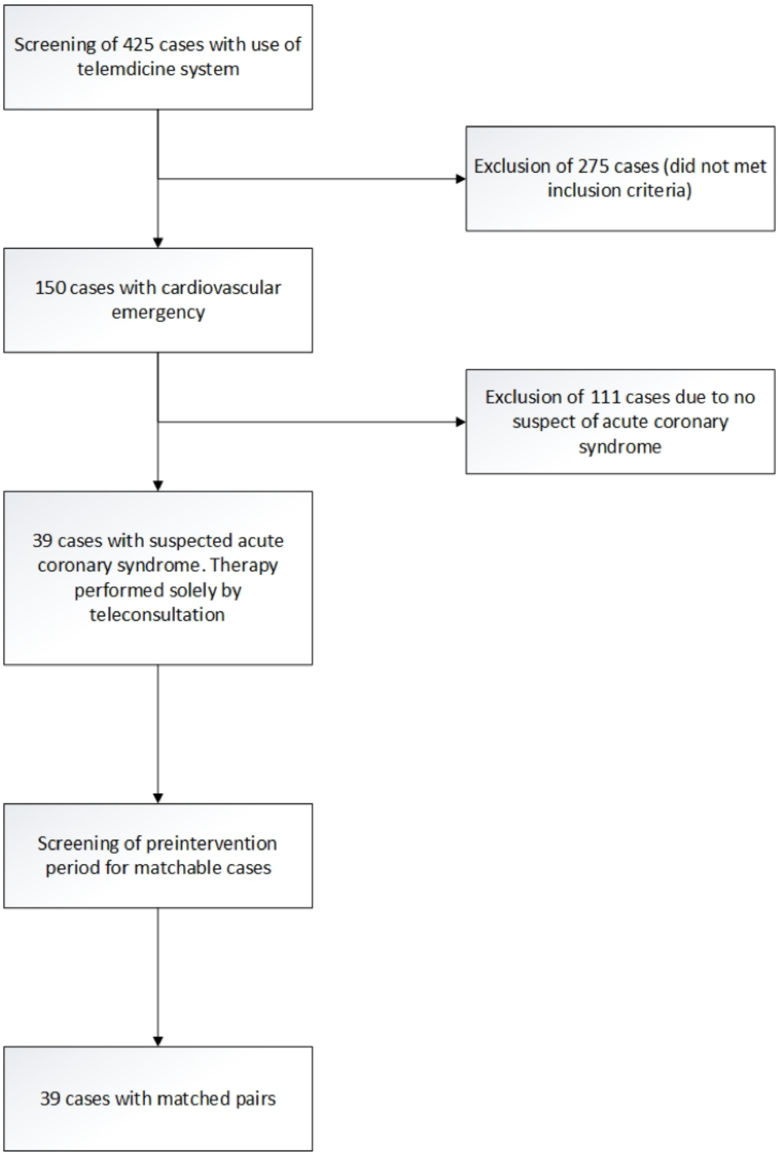
The matched pairs were searched from a historical preintervention period (April 1, 2011 to March 31, 2012) that was one consecutive year prior to the first training lesson. Matching was done using predefined criteria (Table 2) and the nearest-neighbour matching method for all cases that fulfilled the matching criteria.

Table 2. Matching criteria.

Matching criteria	Matching categories
EMS[a] district	Study group cases and control group cases had to be from same district
Date of emergency mission	During the historical control period, the same month and calendar day was used as starting point for the backward and forward search; the first case that fulfilled the matching criteria listed below was included
Patient's age	Same age ± 10 years
Sex	Female
	Male
Symptoms	Chest pain (typically)
	No chest pain (atypically), but other symptoms (ie, nausea, vomiting)
	No symptoms

[a] EMS: emergency medical service.

Figure 1. Study flowchart.



Statistical Methods

Categorical parameters were compared using 4 field contingency tables with McNemar's Test. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc). *P* values <.05 were considered to be significant. Multimedia Appendix 3 shows the statistical evaluation.

Results

Overall, 425 emergency teleconsultation cases were performed with 150 patients having a cardiovascular emergency. The demographic data (Multimedia Appendix 4) was spread equally between the study group and the control group. In each group, 14 patients were female (36%, 14/39), the rest were male. In 39 cases, teleconsultation was started due to suspected ACS. In the described cases, no adverse events were detected. Correct handling of 12-lead ECG was performed equally between the groups (study group, n=38 vs control group, n=39; *P*>.99). There was no significant difference in correct handling of intravenous administration of acetylsalicylic acid (n=38), heparin (n=38), or morphine (n=29) between both groups (study group vs control group): acetylsalicylic acid, n=31 vs n=33, *P*=.73; unfractionated heparin, n=34 vs n=33, *P*>.99; and morphine, n=29 vs n=27 *P*=.50. The correct handling of oxygen differed significantly between the groups (n=34). In the study group it was performed correctly for n=29 and for n=18 in the control group (*P*=.007). Multimedia Appendix 5 gives a detailed view of the correct or incorrect handling of each measure in each of the participating districts. Table 3 summarizes the above mentioned results.

Table 3. Summary of results of correct handling of measures.

Measures	n	Study group	Control group	*P* value
12-lead ECG[a]	39	38	39	>.99
Acetylsalicylic acid	38	31	33	.73
Heparin	38	34	33	>.99
Morphine	29	29	27	.50
Oxygen	34	29	18	.007

[a] ECG: electrocardiogram.

Discussion

Principal Findings

To the best of our knowledge, this is the first study providing prehospital diagnosis and emergency care for ACS that was completely delivered by telemedically supported paramedics. The telemedical concept is feasible, and guideline adherence was at least comparable with on-scene physician care with improved guideline adherence regarding oxygen administration. The feasibility and improvements on outcome by wireless transmission of ECGs to specialists have already been shown in STEMI [4,5,7]. As demonstrated here, both groups handled the administration of the 12-lead ECG correctly. Due to legal restrictions in Germany, the full administration of required medication in patients with ACS administered by paramedics is completely new to the field of prehospital care. The required medication for each patient in the study group was delegated by the tele-EMS physician. Between groups, there were no significant differences in the correct handling of intravenous administration of acetylsalicylic acid, heparin, or morphine. It was already shown that telemedically supported administration of analgesics by paramedics is feasible and secure [22]. Regarding the pain therapy in ACS described here, both groups could improve the correct handling of morphine. We noted that morphine should be administered when NRS was ≥3 (repeatedly 3-5 mg iv). Eventually, morphine was given in both groups even if NRS was not ≥3 or pain reduction was not achieved in a sufficient extent.

In the study group, the oxygen administration was performed significantly more correctly than in the control group. This may be because the tele-EMS physician had a predefined algorithm for correct treatment of ACS at any time in digital form that was integrated in the documentation system. As this algorithm stated, oxygen had to be applied when the blood saturation was ≤95%. We assume that oxygen was applied in the control group all too often, even if the oxygen saturation was above 95%. Therefore, the differences are explained by improved guideline adherence.

Neither the telemedicine group nor the control group had any medical complication. Due to the small sample size, a definitive statement regarding the safety of complete prehospital ACS therapy by paramedics and tele-EMS physicians cannot be given. However, these findings support the thesis that telemedically supported prehospital diagnosis and therapy on the whole is a safe procedure, although the process was completely new for the participating physicians as well as for the paramedics. In comparison with ground- or helicopter-based EMS physician operations [23], this concept allows for a spatially unrestricted emergency concept with fewer resources. At this point it has to be mentioned that, to our knowledge, most studies with 12-lead-ECG transmission represent rather small sample sizes. Therefore, a large randomized controlled trial is needed.

No case had major technical problems. Although the technical performance of the telemedicine system has not yet been evaluated in detail, no consultation had to be terminated due to technical problems.

Limitations

This pilot study had no formal sample size calculation and power analysis. Therefore, all results have to be interpreted against this background. This was the first pilot study to evaluate the concept of telemedically-assisted prehospital therapy in ACS to allow safe, effective, and guideline-adherent therapy. On the basis of these pilot results, a confirmatory trial with a calculated sample size would be meaningful. However, it cannot be definitely stated that telemedically assisted treatment is as effective as treatment with an on-site EMS physician. Due to medical privacy policy in Germany, we could not obtain in-hospital follow-up data and therefore could not highlight outcome parameters such as in-hospital time intervals in STEMI (contact to balloon time, arrival to balloon time), the rate of secondary transfer for PCI (rate of secondary transfer to a different facility for PCI), or laboratory analysis like troponin-levels.

Conclusions

Prehospital diagnosis and treatment of ACS that was carried out by paramedics supported by tele-EMS physicians seems to be as safe and effective as the conventional procedure, assuming a holistic telemedical support system is used, as shown here. In fact, the supply and the adherence to national and international guidelines was better and not worse than in the conventional EMS physician-based service. Further trials with larger patient numbers and a randomized allocation are needed to confirm the findings of this pilot study.

Acknowledgments

The study was conducted within the joint research project “TemRas,” funded by the European Union and the Ministry of Innovation, Science and Research of North Rhine Westphalia, Germany (MIWF), Project-No: PtJ-Az. 0909im002b, Funding code 005-1003-0034. Philips Healthcare (Hamburg, Germany) and P3 communications (Aachen, Germany) contributed their own financial resources. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. No author had financial relationships with or other nonfinancial dependencies on the funding sponsors.

The authors thank their colleagues from the Institute of Information Management in Mechanical Engineering (IMA), Centre for Learning and Knowledge Management (ZLW), Associate Institute for Management Cybernetics eV (IfU), ZLW/IMA, and the P3 communications who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations and conclusions of this article.

Authors' Contributions

JCB, FH, RR, and SB made substantial contributions to conception and design, acquisition of data, and drafting the manuscript. MC, CC, and SKB participated in study design. CC and FH acquired data. FH, CC, JB, and SB analyzed and interpreted data. MT carried out statistical analysis of data. All authors were involved in drafting the manuscript and revising it critically. All authors read and approved the final manuscript.

Conflicts of Interest

After the completion of the research project JCB, RR, and MC founded the Docs in Clouds telemedicine service and consulting company.

Multimedia Appendix 1

Trial protocol.

[PDF File (Adobe PDF File), 881KB - jmir_v18i12e314_app1.pdf]

Multimedia Appendix 2

Dataset.

[XLSX File (Microsoft Excel File), 15KB - jmir_v18i12e314_app2.xlsx]

Multimedia Appendix 3

Statistical evaluation (raw dataset).

[PDF File (Adobe PDF File), 71KB - jmir_v18i12e314_app3.pdf]

Multimedia Appendix 4

Demographic data.

[PDF File (Adobe PDF File), 11KB - jmir_v18i12e314_app4.pdf]

Multimedia Appendix 5

Detail table.

[PDF File (Adobe PDF File), 293KB - jmir_v18i12e314_app5.pdf]

References

1. Löwel H, Meisinger C, Heier M, Hörmann A, Kuch B, Gostomzyk J, et al. Sex specific trends of sudden cardiac death and acute myocardial infarction: results of the population-based KORA/MONICA-Augsburg register 1985 to 1998. Dtsch Med Wochenschr 2002 Nov 2;127(44):2311-2316. [doi: 10.1055/s-2002-35181] [Medline: 12410433]
2. Goodman SG, Huang W, Yan AT, Budaj A, Kennelly BM, Gore JM, Expanded Global Registry of Acute Coronary Events (GRACE2) Investigators. The expanded Global Registry of Acute Coronary Events: baseline characteristics, management practices, and hospital outcomes of patients with acute coronary syndromes. Am Heart J 2009 Aug;158(2):193-201.e1. [doi: 10.1016/j.ahj.2009.06.003] [Medline: 19619694]
3. Fox KA, Cokkinos DV, Deckers J, Keil U, Maggioni A, Steg G. The ENACT study: a pan-European survey of acute coronary syndromes. European Network for Acute Coronary Treatment. Eur Heart J 2000 Sep;21(17):1440-1449 [FREE Full text] [doi: 10.1053/euhj.2000.2185] [Medline: 10952836]

4. Adams GL, Campbell PT, Adams JM, Strauss DG, Wall K, Patterson J, et al. Effectiveness of prehospital wireless transmission of electrocardiograms to a cardiologist via hand-held device for patients with acute myocardial infarction (from the Timely Intervention in Myocardial Emergency, NorthEast Experience [TIME-NE]). Am J Cardiol 2006 Nov 1;98(9):1160-1164. [doi: 10.1016/j.amjcard.2006.05.042] [Medline: 17056318]
5. Dhruva VN, Abdelhadi SI, Anis A, Gluckman W, Hom D, Dougan W, et al. ST-Segment analysis using wireless technology in acute myocardial infarction (STAT-MI) trial. J Am Coll Cardiol 2007 Aug 7;50(6):509-513 [FREE Full text] [doi: 10.1016/j.jacc.2007.04.049] [Medline: 17678733]
6. Sejersten M, Sillesen M, Hansen PR, Nielsen SL, Nielsen H, Trautner S, et al. Effect on treatment delay of prehospital teletransmission of 12-lead electrocardiogram to a cardiologist for immediate triage and direct referral of patients with ST-segment elevation acute myocardial infarction to primary percutaneous coronary intervention. Am J Cardiol 2008 Apr 1;101(7):941-946. [doi: 10.1016/j.amjcard.2007.11.038] [Medline: 18359312]
7. Sanchez-Ross M, Oghlakian G, Maher J, Patel B, Mazza V, Hom D, et al. The STAT-MI (ST-Segment analysis using wireless technology in acute myocardial infarction) trial improves outcomes. JACC Cardiovasc Interv 2011 Feb;4(2):222-227 [FREE Full text] [doi: 10.1016/j.jcin.2010.11.007] [Medline: 21349462]
8. Forberg JL, Khoshnood A, Green M, Ohlsson M, Björk J, Jovinge S, et al. An artificial neural network to safely reduce the number of ambulance ECGs transmitted for physician assessment in a system with prehospital detection of ST elevation myocardial infarction. Scand J Trauma Resusc Emerg Med 2012;20:8 [FREE Full text] [doi: 10.1186/1757-7241-20-8] [Medline: 22296816]
9. Reimann B, Maier B, Lott R, Konrad F. Gefaehrdung der Notarztversorgung im laendlichen Gebiet. Notfall & Rettungsmedizin 2004 May;7(3):200-204. [doi: 10.1007/s10049-004-0656-8]
10. Behrendt H, Schmiedel R. Die aktuellen Leistungen des Rettungsdienstes in der Bundesrepublik Deutschland im zeitlichen Vergleich (Teil II). Notfall & Rettungsmedizin 2004 Feb 1;7(1):59-70. [doi: 10.1007/s10049-003-0623-9]
11. Skorning M, Bergrath S, Rörtgen D, Beckers SK, Brokmann JC, Gillmann B, et al. Teleconsultation in pre-hospital emergency medical services: real-time telemedical support in a prospective controlled simulation study. Resuscitation 2012 May;83(5):626-632. [doi: 10.1016/j.resuscitation.2011.10.029] [Medline: 22115932]
12. Rörtgen D, Bergrath S, Rossaint R, Beckers SK, Fischermann H, Na I, et al. Comparison of physician staffed emergency teams with paramedic teams assisted by telemedicine--a randomized, controlled simulation study. Resuscitation 2013 Jan;84(1):85-92. [doi: 10.1016/j.resuscitation.2012.06.012] [Medline: 22750663]
13. Bergrath S, Rörtgen D, Rossaint R, Beckers SK, Fischermann H, Brokmann JC, et al. Technical and organisational feasibility of a multifunctional telemedicine system in an emergency medical service - an observational study. J Telemed Telecare 2011;17(7):371-377. [doi: 10.1258/jtt.2011.110203] [Medline: 21933897]
14. Bergrath S, Reich A, Rossaint R, Rörtgen D, Gerber J, Fischermann H, et al. Feasibility of prehospital teleconsultation in acute stroke--a pilot study in clinical routine. PLoS One 2012;7(5):e36796 [FREE Full text] [doi: 10.1371/journal.pone.0036796] [Medline: 22629331]
15. Bergrath S, Czaplik M, Rossaint R, Hirsch F, Beckers SK, Valentin B, et al. Implementation phase of a multicentre prehospital telemedicine system to support paramedics: feasibility and possible limitations. Scand J Trauma Resusc Emerg Med 2013;21:54 [FREE Full text] [doi: 10.1186/1757-7241-21-54] [Medline: 23844941]
16. Czaplik M, Bergrath S, Rossaint R, Thelen S, Brodziak T, Valentin B, et al. Employment of telemedicine in emergency medicine. Clinical requirement analysis, system development and first test results. Methods Inf Med 2014;53(2):99-107. [doi: 10.3414/ME13-01-0022] [Medline: 24477815]
17. Hamm CW. Guidelines: Acute coronary syndrome (ACS). II: Acute coronary syndrome with ST-elevation. Z Kardiol 2004 Apr;93(4):324-341. [doi: 10.1007/s00392-004-0109-x] [Medline: 15085379]
18. Hamm CW. Guidelines: acute coronary syndrome (ACS). 1: ACS without persistent ST segment elevations. Z Kardiol 2004 Jan;93(1):72-90. [doi: 10.1007/s00392-004-1064-2] [Medline: 14740245]
19. Task Force on the management of ST-segment elevation acute myocardial infarction of the European Society of Cardiology (ESC), Steg PG, James SK, Atar D, Badano LP, Blömstrom-Lundqvist C, et al. ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J 2012 Oct;33(20):2569-2619 [FREE Full text] [doi: 10.1093/eurheartj/ehs215] [Medline: 22922416]
20. Hamm CW, Bassand JP, Agewall S, Bax J, Boersma E, Bueno H, ESC Committee for Practice Guidelines. ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force for the management of acute coronary syndromes (ACS) in patients presenting without persistent ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J 2011 Dec;32(23):2999-3054 [FREE Full text] [doi: 10.1093/eurheartj/ehr236] [Medline: 21873419]
21. Antman EM, Anbe DT, Armstrong PW, Bates ER, Green LA, Hand M, American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 1999 Guidelines for the Management of Patients With Acute Myocardial Infarction). ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction--executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 1999 Guidelines for the Management of Patients With

Acute Myocardial Infarction). Circulation 2004 Aug 3;110(5):588-636 [FREE Full text] [doi: 10.1161/01.CIR.0000134791.68010.FA] [Medline: 15289388]

22. Brokmann JC, Rossaint R, Hirsch F, Beckers SK, Czaplik M, Chowanetz M, et al. Analgesia by telemedically supported paramedics compared with physician-administered analgesia: A prospective, interventional, multicentre trial. Eur J Pain 2016 Aug;20(7):1176-1184. [doi: 10.1002/ejp.843] [Medline: 26914284]
23. Schuster M, Pints M, Fiege M. Duration of mission time in prehospital emergency medicine: effects of emergency severity and physicians level of education. Emerg Med J 2010 May;27(5):398-403. [doi: 10.1136/emj.2009.074211] [Medline: 20442179]

Abbreviations

ACS: acute coronary syndrome
ECG: electrocardiogram
EMS: emergency medical service
iv: intravenous
IU: international units
NRS: numeric rating scale
NSTE-ACS: non–STEMI-acute coronary syndrome
PCI: percutaneous coronary intervention
STEMI: ST elevation myocardial infarction
TemRas: Telemedical rescue assistance system

Edited by G Eysenbach; submitted 15.07.16; peer-reviewed by S Omboni, N Bashi; comments to author 25.08.16; revised version received 10.10.16; accepted 15.10.16; published 01.12.16.

Please cite as:
Brokmann JC, Conrad C, Rossaint R, Bergrath S, Beckers SK, Tamm M, Czaplik M, Hirsch F
Treatment of Acute Coronary Syndrome by Telemedically Supported Paramedics Compared With Physician-Based Treatment: A Prospective, Interventional, Multicenter Trial
J Med Internet Res 2016;18(12):e314
URL: http://www.jmir.org/2016/12/e314/
doi:10.2196/jmir.6358
PMID:27908843

Original Paper

Recruiting Pregnant Patients for Survey Research: A Head to Head Comparison of Social Media-Based Versus Clinic-Based Approaches

Lindsay Admon[1,2], MD; Jessica K Haefner[2,3], BSc; Giselle E Kolenic[2], MA; Tammy Chang[3,4], MPH, MS, MD; Matthew M Davis[5], MAPP, MD; Michelle H Moniz[2,3], MSc, MD

[1]National Clinician Scholars Program, University of Michigan Institute for Healthcare Policy and Innovation and the US Department of Veterans Affairs, Ann Arbor, MI, United States

[2]Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, MI, United States

[3]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States

[4]Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States

[5]Department of Pediatrics, Northwestern Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Lindsay Admon, MD
National Clinician Scholars Program
University of Michigan Institute for Healthcare Policy and Innovation and the US Department of Veterans Affairs
North Campus Research Complex
2800 Plymouth Road, Bldg 14 Room G100-22
Ann Arbor, MI, 48109-2800
United States
Phone: 1 734 647 0571
Fax: 1 734 647 3301
Email: lindskb@med.umich.edu

Abstract

Background: Recruiting a diverse sample of pregnant women for clinical research is a challenging but crucial task for improving obstetric services and maternal and child health outcomes.

Objective: To compare the feasibility and cost of recruiting pregnant women for survey research using social media-based and clinic-based approaches.

Methods: Advertisements were used to recruit pregnant women from the social media website Facebook. In-person methods were used to recruit pregnant women from the outpatient clinic of a large, tertiary care center. In both approaches, potential respondents were invited to participate in a 15-minute Web-based survey. Each recruitment method was monitored for 1 month. Using bivariate statistics, we compared the number, demographic characteristics, and health characteristics of women recruited and the cost per completed survey for each recruitment method.

Results: The social media-based approach recruited 1178 women and the clinic-based approach recruited 219 women. A higher proportion of subjects recruited through social media identified as African American (29.4%, 207/705 vs 11.2%, 20/179), reported household incomes <US $30,000 per year (56.8%, 409/720 vs 25.8%, 47/182), reported being in early pregnancy (18.6%, 135/726 vs 10.4%, 19/183 first trimester), and rated their health as fair or poor (22.2%, 160/722 vs 8.2%, 15/183; all *P*<.001). A smaller proportion of subjects recruited through social media had earned a college degree (21.3%, 153/717 vs 62.3%, 114/183) and were married or in a domestic partnership (45.7%, 330/722 vs 72.1%, 132/183; all *P*<.001). Social media-based recruitment costs were US $14.63 per completed survey, compared with US $23.51 for clinic-based recruitment.

Conclusions: Web-based recruitment through a social networking platform is a feasible, inexpensive, and rapid means of recruiting a large, diverse sample of pregnant women for survey research.

(J Med Internet Res 2016;18(12):e326) doi:10.2196/jmir.6593

KEYWORDS

pregnant women; surveys and questionnaires; methods; social media

Introduction

Recruiting pregnant women for clinical research is a challenging but crucial task for improving obstetric services and maternal and child health outcomes. Conventional methods for recruiting pregnant women for survey research face significant limitations. Traditionally, pregnant women have been recruited from clinical care sites, an approach that is often plagued by underrepresentation of women in early pregnancy, poor demographic diversity, and the inability to access women who do not seek prenatal care—resulting in limited generalizability of study findings and inferential errors and bias in the use of such data [1,2]. Recruitment through established national samples (eg, Survey Sampling International, Gesellschaft fur Konsumforschung or GfK) is attractive for its methodological rigor and geographic diversity of participants. However, these companies are often only able to provide small samples of pregnant women, which limit analytic power. As a result, innovative strategies are needed to conduct methodologically sound and cost-effective survey research influencing the care of pregnant women.

The Internet and social media platforms offer promising avenues for recruitment of pregnant women. Recent data from the Pew Internet & American Life Project reveal that 96% of 18- to 29-year-olds in the United States have access to the Internet and that Internet use continues to rise among the adult population [3]. An even more recent phenomenon is the rise in popularity of social networking websites, with Facebook (Facebook Inc., Palo Alto, CA, USA) second only to Google (Google Inc., Mountain View, CA, USA) as the most popular website in the United States [4]. With 1.09 billion active users daily, Facebook has been used as an effective method to recruit female adolescents and women for survey research [1,5-15]. Little is known, however, about whether social media platforms are a feasible method for recruitment of pregnant women and how they compare with traditional recruitment methods in terms of cost.

Our team conducted a head-to-head comparison of 2 recruitment methods among pregnant women: Web-based recruitment through social media and traditional in-person recruitment at prenatal clinics. Our objective was to describe the relative advantages and disadvantages of each approach, including the total number of women recruited, their demographic and health characteristics, and the cost per completed survey for each recruitment method.

Methods

Recruitment

We utilized a cross-sectional design to invite respondents recruited via both methods to watch a brief Web-based video about health in pregnancy and answer survey items about pregnancy-related health knowledge and behaviors. Approximately 1 month of time was allocated to each recruitment method. Women were eligible for participation if they self-identified as being 18 years of age or older, pregnant, English speaking, and living in the United States. The study was deemed exempt by the study site's institutional review board.

Recruitment of the Social Media-Based Sample

We recruited participants through advertisements on Facebook. Ads created on Facebook contained three key features: an image, a caption, and an ad copy containing a link to the survey website. Study ads were developed by the investigators using Facebook's self-service application. Ads were shown on Facebook users' newsfeeds. Clicking on the hyperlink within the ad led users to our survey website.

Facebook's platform can direct ads to specific audiences based on sex, age, location, and interests. Interest-based targeting parameters such as "expectant parents," "interest in motherhood," and "parents"—or a combination of these—may be utilized. Facebook uses a proprietary method to infer users' racial and ethnic "interests" by analyzing pages and posts users have liked or engaged with on Facebook. These inferences can be used as targeting parameters. Ad content can also be tailored. For example, ad images may differ in the race of women depicted.

Before study recruitment began, ads were refined over 11 days during May 2015 by testing a variety of combinations of image, caption, ad copy, and interest-based targeting parameters. During this ad refinement phase, Facebook automatically monitored the cost per click on the survey link included in each potential recruitment ad, as well as the cost per completed survey. This enabled the study team to learn which ads worked best for progressive outreach to audiences of particular interest. To investigate the effects of financial incentives on recruitment, US $5 and US $10 incentives were tested during this period, with results showing that the US $10 offer attracted more clicks. Availability of incentives was indicated in the ads that appeared on Facebook. Respondents who completed the survey during the ad refinement period, 21 in total, are not included in the recruitment numbers.

The final campaign included the highest-performing ads with US $10 incentives, ultimately targeting women aged 18 years and older, living in the United States, and fitting the following interest profiles: African American expectant parents, Asian American expectant parents, and Hispanic women interested in motherhood. A sample Facebook ad is provided in Figure 1.

Figure 1. Sample Facebook advertisement used for social media–based recruitment of pregnant women.

Living for Two

Sponsored ·

Participate in a research study by enjoying a brief video and taking a quick survey. Get $10 now.

Baby on the way?

umichumhs.qualtrics.com

Following pilot testing, women were recruited through Facebook during 27 consecutive days in May-June 2015 utilizing the targeting parameters listed above. Potential participants were able to engage with the study materials via their own Internet access (mobile devices, personal computers, public computers, etc).

Recruitment of the Clinical Sample

Clinic-based recruitment occurred over 29 business days in August-September 2015. Pregnant women were recruited at routine outpatient obstetric visits at a large tertiary care center in the Midwest. Clinic clerks distributed study fliers to all patients presenting for prenatal appointments. Patients who expressed interest in participating were directed to trained research staff in the waiting room. Research staff invited participation, assessed eligibility, and answered questions about the study based on a standardized recruitment script. Patients completed the survey in the clinic waiting room or examination room using a tablet and headphones provided by the study team. Research staff sent participants individual survey links by email if they preferred to take the survey via their own Internet access. No identifying information was collected, and survey responses were not linked to participants' clinical information. Because of the high proportion of women interested in participation with the US $5 incentive, US $10 incentives were not offered.

Survey Completion

The survey's introductory webpage explained the purpose of the study, the anonymous nature of the research, the expected time needed to complete the survey (15 minutes), and the option to quit at any time. Upon completion of the survey, participants were asked to provide their email address for delivery of the incentive. Participant email addresses were reviewed for repetition or similarity to exclude duplicate respondents who appeared to be seeking multiple incentives.

Statistical Analysis

All data were deidentified and analyzed in Stata 14 (StataCorp LP).

Demographics

Demographic information and health characteristics were first investigated with descriptive statistics, including means and proportions, and stratified by recruitment method. Independent samples *t* tests and chi-square tests of independence were used to investigate whether or not these measures differed significantly by recruitment method.

Costs

The direct research-related cost per completed survey was calculated for each recruitment method. For social media-based recruitment, cost per completed survey was calculated as follows: ([pilot ad cost + pilot-testing respondent incentives] + [recruitment ad cost + respondent incentives])/number of completed surveys as a result of Facebook approaches. For the clinical sample, cost per completed survey was calculated as follows: (respondent incentives + research assistants' salaries)/number of completed surveys as a result of clinic approaches. Investigator costs were not included as the study team developed and deployed the campaigns as a group and focused on the Web-based and clinical efforts in a consistent fashion across both modes.

Results

Recruitment

A flow diagram comparing social media-based with clinical recruitment is presented in Figure 2. Facebook ads were shown on 364,035 users' newsfeeds over the 4-week campaign period. There were 9972 clicks on the ads, resulting in 1323 entries to the survey's webpage and 1178 respondents who consented to participate. Among consenting respondents recruited via

Facebook, 74.02% (872/1178) met eligibility criteria and 64.43% (759/1178) completed the survey.

During in-person clinical recruitment, approximately 500 unique pregnant patients were seen in clinic. Among the patients consenting to participate, 95.9% (210/219) met eligibility criteria and 190 86.8% (190/219) completed the survey.

Demographics

A higher proportion of subjects recruited through social media self-identified as African American (29.4%, 207/705 vs 11.2%, 20/179; *P*<.001; Table 1) and reported annual household incomes <US $30,000 per year (56.8%, 409/720 vs 25.8%, 47/182; *P*<.001). A significantly lower proportion of those recruited through social media had earned a college degree (21.3%, 153/717 vs 62.3%, 114/183; *P*<.001) and were married or in a domestic partnership (45.7%, 330/722 vs 72.1%, 132/183; *P*<.001). With respect to health status, a higher proportion of those recruited through social media were in the first trimester of pregnancy (18.6%, 135/726 vs 10.4%, 19/183; *P*<.001) and rated their own health as fair or poor with greater frequency (22.2%, 160/722 vs 8.2%, 15/183; *P*<.001).

Cost

The total cost of social media-based recruitment was calculated by adding the pilot-testing cost (pilot ads US $494.51 + pilot incentives US $155.00) and the final campaign cost (ads US $3243.74 + incentives US $7210.00), which totaled US $11,103.25 (Figure 3). Dividing the total cost by the total number of surveys completed as a result of Facebook approaches (n=759), the cost per completed survey was determined to be US $14.63.

The total cost of recruiting the clinical sample was calculated by adding expenditures on research assistants' salaries (US $3551.68) and on participant incentives (US $915.00), which totaled US $4466.68. Dividing the total cost by the total number of surveys completed as a result of in-clinic approaches (n=190), the cost per completed survey was determined to be US $23.51.

Figure 2. Flow diagram of inclusion, exclusion, and dropout of pregnant women recruited via social media–based compared to clinic-based approaches.

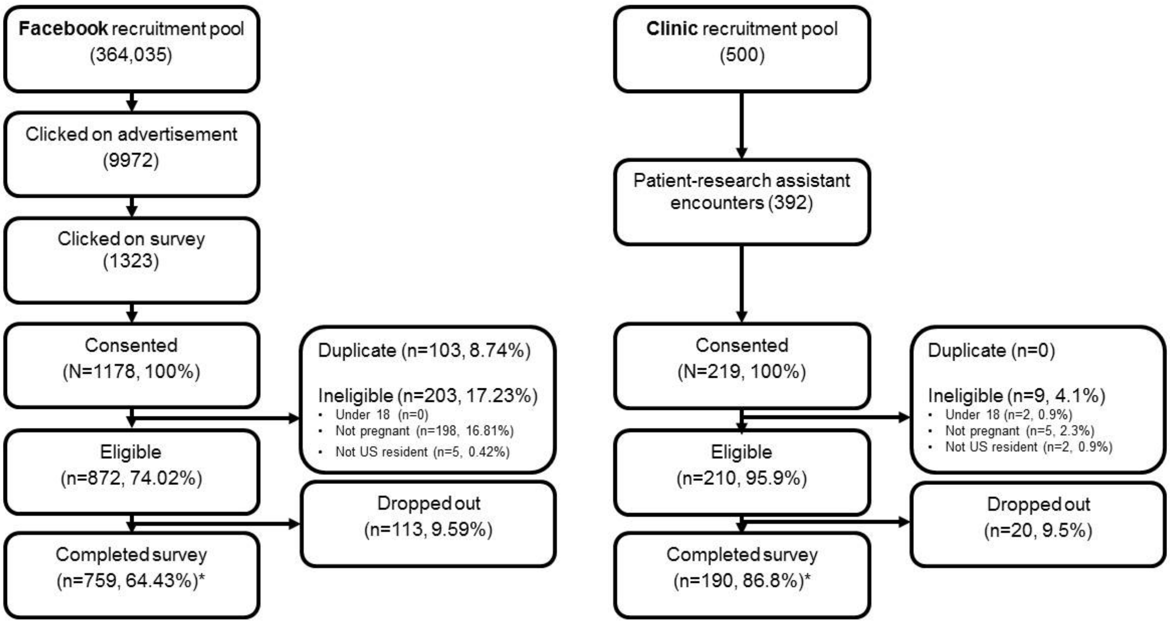


Table 1. Characteristics of pregnant women recruited via social media-based versus clinic-based approaches.

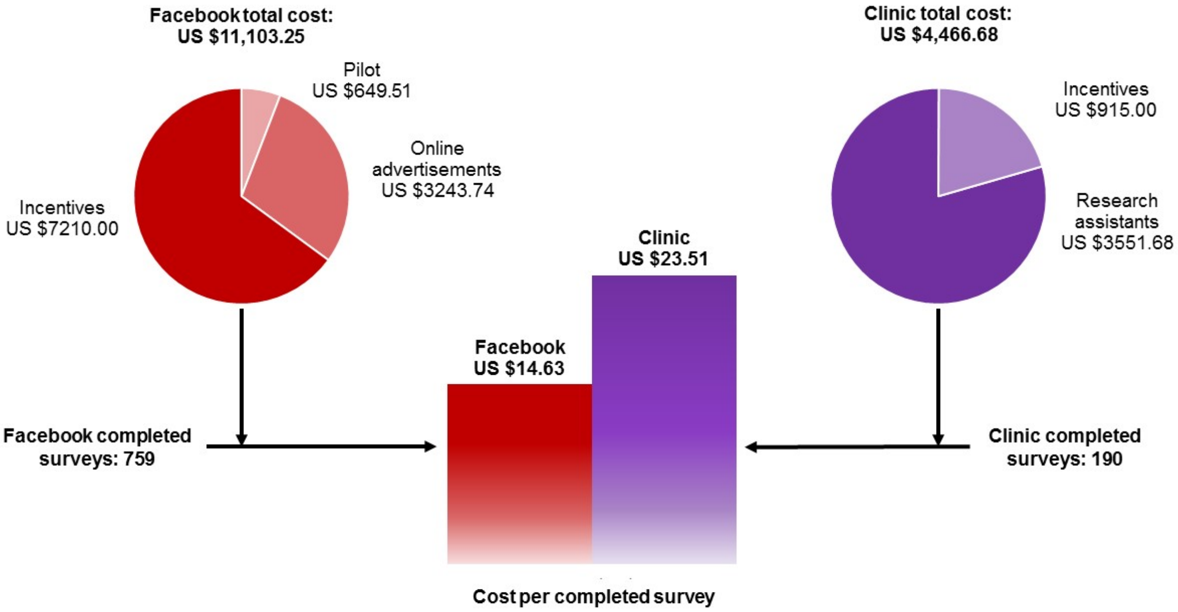
Variable	Social media-based recruitment (n=759)	Clinic-based recruitment (n=190)	*P* value
Age in years, mean (SD)[a]	27.3 (5.2)	29.9 (4.7)	<.001
Gestational age, n (%)	n=726	n=183	
<14 weeks	135 (18.6)	19 (10.4)	
14-27 weeks	276 (38.0)	54 (29.5)	
>28 weeks	315 (43.4)	110 (60.1)	<.001
Race, n (%)	n=705	n=179	
White only	283 (40.1)	130 (72.6)	
African American only	207 (29.4)	20 (11.2)	
Multiracial or other	215 (30.5)	29 (16.2)	<.001
Education, n (%)	n=717	n=183[b]	
Did not complete high school	42 (5.9)	5 (2.7)	
High school diploma	220 (30.7)	20 (10.9)	
Associate’s degree or some college	302 (42.1)	44 (24.0)	
College graduate	153 (21.3)	114 (62.3)	<.001
Relationship status, n (%)	n=722	n=183	
Single	73 (10.1)	8 (4.4)	
In a relationship	319 (44.2)	43 (23.5)	
Married or domestic partnership	330 (45.7)	132 (72.1)	<.001
Household income, US $, n (%)	n=720	n=182	
<30,000	409 (56.8)	47 (25.8)	
30,000-60,000	231 (32.1)	37 (20.3)	
>60,000	80 (11.1)	93 (53.9)	<.001
Rating of own health, n (%)	n=722	n=183	
Excellent or very good	276 (38.2)	106 (57.9)	
Good	286 (39.6)	62 (33.9)	
Fair or poor	160 (22.2)	15 (8.2)	<.001
Region[c], n (%)	n=563[b]		
Northeast	32 (5.7)		
Midwest	342 (60.8)		
South	114 (20.3)		
West	75 (13.3)		

[a] Social media-based recruitment n=750.

[c] Sum of percentages for subpopulation is within 0.1 of 100.0 due to rounding error.

[c] The region was not asked in clinic-based recruitment as the participants were recruited from a clinic in the Midwest.

Figure 3. Cost per completed survey among pregnant women recruited via social media–based compared to clinic-based approaches.



Discussion

Principal Findings

Our findings indicate that Web-based recruitment may be a feasible approach to recruit a large, diverse sample of pregnant women at low cost relative to traditional, clinic-based recruitment. Over similar periods of time, the social media-based approach generated 5 times as many pregnant women for survey research compared to clinic recruitment. Women recruited through social media were also more demographically diverse than clinic recruits, and social media-based recruitment cost less per completed survey. Given that participants reported early gestational ages and were widely distributed geographically, it is unlikely that a similar cohort could have been recruited in 1 month using traditional in-clinic methods, even at a higher cost.

Recruitment of racially and ethnically diverse samples of pregnant women for research has been identified as a priority by the Division of Reproductive Health at the Centers for Disease Control and Prevention [16]. Adverse pregnancy outcomes, including severe maternal morbidity and mortality, are disproportionally higher among many minority populations [16,17]. At the same time, racially and ethnically diverse populations have been historically underrepresented in clinical research efforts for reasons including failure on behalf of researchers to reach minorities [18]. Furthermore, the association between socioeconomic disadvantage and poorer maternal and fetal pregnancy outcomes has long been established [19,20], as has the large and persistent association between lower educational attainment and worse health [21]. Our findings support the feasibility of social media to access vulnerable populations who may be harder to reach through traditional, clinic-based approaches.

Lower costs for social media-based recruitment (US $14.63 per completed survey) versus clinic-based recruitment (US $23.51 per completed survey), in combination with faster recruitment over time, suggest that social media-based recruitment offers a so-called dominant strategy—the process is more favorable and costs less. Importantly, it is possible that recruitment goals for either method could be met in shorter time with larger incentives or at cheaper cost over a longer period. Post hoc consideration of cost data for clinic recruitment revealed that the *rate* of accruing new subjects decreased over time, while the *cost* of accruing new subjects increased over time—indicating decreasing efficiency. In contrast, likely due to the exponentially larger pool of eligible participants on Facebook, the rate of subject accrual and cost were linear over time, which may make projections of recruitment through social media more predictable. While time-to-saturation of clinical recruitment pools may vary based on the characteristics of different patient populations, our findings may guide other researchers in estimating their own costs for reaching recruitment goals.

Despite these advantages, recruitment through social media has important disadvantages. This recruitment approach involves snowball sampling within established social networks; therefore, it may not be ideal for surveys trying to measure diversity of attitudes, perceptions, or experiences in care. For instance, snowball sampling likely explains the overrepresentation of those from the Midwest in the Facebook recruitment population. Fortunately, an inherent capability of the Facebook approach is that ad performance and cost can be monitored and adjusted in real time, which requires minimal time or technical expertise. Investigators can engage in purposive sampling by targeting ads to underrepresented users based on changing targets for demographic characteristics or interests as the study progresses.

Limitations

Limitations of the study design should be considered when interpreting our findings. While clinic-based recruitment may have produced a more diverse sample in different or more numerous clinical settings, recruited participants will largely reflect the sample of patients served in a specific clinical setting. The Internet, in contrast, offers an inherently broader, more diverse recruitment pool and the potential advantages of targeting specific individuals or groups. Next, eligibility was assessed by self-report. It is unclear whether reliability of self-report differs between Web-based versus in-person recruitment settings, and comparative assessment of the quality of data obtained through social media is an important direction for future research [22]. Because of failure to meet eligibility criteria, there was a notably higher dropout rate in the Facebook sample. In future work, Facebook ads could be edited to more clearly reflect requirements for participation. Finally, there were also some indirect research costs, which were unclear in terms of their implications and are not included in reported estimates. These include 6 phone calls with Facebook and the labor required to train clinic clerks and research assistants, which required approximately 10 hours of the authors' time in total and were split across the recruitment modalities.

Conclusions

Successful survey research designed to improve the care of pregnant women requires the recruitment of diverse, adequately powered samples of participants. There are significant barriers to achieving this outcome with traditional, in-person recruitment, and the extant literature offers little comparative guidance regarding recruitment methods. Our head-to-head comparison of social media-based versus clinic-based recruitment of pregnant patients for survey research suggests that the use of a social networking platform is a feasible, inexpensive, and efficient approach to recruiting a large, diverse sample of pregnant women for survey research.

Acknowledgments

We thank Cosette Kathawa and Morgen Barroso for their assistance in the clinical recruitment phase of this project. This study was funded by the University of Michigan Health System and by a grant from the state of Michigan's Department of Health and Human Services; neither organization played any role in study design, analysis and interpretation of data, writing of the report, or the decision to submit the manuscript for publication. During the study period, LA was supported by the National Clinician Scholars Program at the University of Michigan Institute for Healthcare Policy and Innovation and the US Department of Veterans Affairs, MHM was supported by the Robert Wood Johnson Foundation Clinical Scholars Program, and MMD served as the chief medical executive for the state of Michigan, none of which were involved in this study.

Conflicts of Interest

None declared.

References

1. Arcia A. Facebook advertisements for inexpensive participant recruitment among women in early pregnancy. Health Educ Behav 2013 Sep 30;41(3):237-241 [FREE Full text] [doi: 10.1177/1090198113504414] [Medline: 24082026]
2. Gatny HH, Axinn WG. Willingness to participate in research during pregnancy: race, experience, and motivation. Field methods 2011 Oct 9;24(2):135-154 [FREE Full text] [doi: 10.1177/1525822X11419819] [Medline: 22798727]
3. Perrin A, Duggan M. Americans' Internet Access: 2000-2015. 2015 Jun 26. Pew Research Center. URL: http://www.pewinternet.org/2015/06/26/americans-internet-access-2000-2015/#internet-usage-by-age [accessed 2016-12-12] [WebCite Cache ID 6kB6lXCtU]
4. Alexa. How popular is facebook.com? URL: http://www.alexa.com/siteinfo/facebook.com [accessed 2016-08-31] [WebCite Cache ID 6mi53V3Jy]
5. Amon KL, Campbell AJ, Hawke C, Steinbeck K. Facebook as a recruitment tool for adolescent health research: a systematic review. Acad Pediatr 2014;14(5):439-447.e4. [doi: 10.1016/j.acap.2014.05.049] [Medline: 25169155]
6. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
7. Jones L, Saksvig BI, Grieser M, Young DR. Recruiting adolescent girls into a follow-up study: benefits of using a social networking website. Contemp Clin Trials 2012 Mar;33(2):268-272 [FREE Full text] [doi: 10.1016/j.cct.2011.10.011] [Medline: 22101207]
8. Battistella E, Kalyan S, Prior JC. Evaluation of methods and costs associated with recruiting healthy women volunteers to a study of ovulation. J Womens Health (Larchmt) 2010 Aug;19(8):1519-1524. [doi: 10.1089/jwh.2009.1751] [Medline: 20626268]
9. Harris ML, Loxton D, Wigginton B, Lucke JC. Recruiting online: lessons from a longitudinal survey of contraception and pregnancy intentions of young Australian women. Am J Epidemiol 2015 May 15;181(10):737-746. [doi: 10.1093/aje/kwv006] [Medline: 25883155]

10. Lord S, Brevard J, Budman S. Connecting to young adults: an online social network survey of beliefs and attitudes associated with prescription opioid misuse among college students. Subst Use Misuse 2011;46(1):66-76 [FREE Full text] [doi: 10.3109/10826084.2011.521371] [Medline: 21190407]
11. Chu JL, Snider CE. Use of a social networking web site for recruiting Canadian youth for medical research. J Adolesc Health 2013 Jun;52(6):792-794. [doi: 10.1016/j.jadohealth.2012.12.002] [Medline: 23352727]
12. Williams MT, Proetto D, Casiano D, Franklin ME. Recruitment of a hidden population: African Americans with obsessive-compulsive disorder. Contemp Clin Trials 2012 Jan;33(1):67-75 [FREE Full text] [doi: 10.1016/j.cct.2011.09.001] [Medline: 21983626]
13. 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]
14. Facebook. Our Mission. URL: https://newsroom.fb.com/company-info/ [accessed 2016-12-12] [WebCite Cache ID 6kB7R4Fbj]
15. Adam LM, Manca DP, Bell RC. Can Facebook be used for research? Experiences using Facebook to recruit pregnant women for a randomized controlled trial. J Med Internet Res 2016;18(9):e250 [FREE Full text] [doi: 10.2196/jmir.6404] [Medline: 27655184]
16. Creanga AA, Berg CJ, Syverson C, Seed K, Bruce FC, Callaghan WM. Pregnancy-related mortality in the United States, 2006-2010. Obstet Gynecol 2015 Jan;125(1):5-12. [doi: 10.1097/AOG.0000000000000564] [Medline: 25560097]
17. Xu J, Kochanek KD, Murphy SL, Tejada-Vera B. Deaths: final data for 2007. Natl Vital Stat Rep 2010 May;58(19):1-19 [FREE Full text] [Medline: 25075874]
18. Shavers-Hornaday VL, Lynch CF, Burmeister LF, Torner JC. Why are African Americans under-represented in medical research studies? Impediments to participation. Ethn Health 1997;2(1-2):31-45. [doi: 10.1080/13557858.1997.9961813] [Medline: 9395587]
19. Blumenshine P, Egerter S, Barclay CJ, Cubbin C, Braveman PA. Socioeconomic disparities in adverse birth outcomes: a systematic review. Am J Prev Med 2010 Sep;39(3):263-272. [doi: 10.1016/j.amepre.2010.05.012] [Medline: 20709259]
20. Bryant AS, Worjoloh A, Caughey AB, Washington AE. Racial/ethnic disparities in obstetric outcomes and care: prevalence and determinants. Am J Obstet Gynecol 2010 Apr;202(4):335-343 [FREE Full text] [doi: 10.1016/j.ajog.2009.10.864] [Medline: 20060513]
21. Cutler D, Lleras-Muney A. NPC Policy Brief #9. National Poverty Center. 2007 Mar. URL: http://www.npc.umich.edu/publications/policy_briefs/brief9/policy_brief9.pdf [accessed 2016-12-12] [WebCite Cache ID 6kB7ZWc5a]
22. Alshaikh F, Ramzan F, Rawaf S, Majeed A. Social network sites as a mode to collect health data: a systematic review. J Med Internet Res 2014;16(7):e171 [FREE Full text] [doi: 10.2196/jmir.3050] [Medline: 25048247]

Edited by G Eysenbach; submitted 08.09.16; peer-reviewed by A Arcia, C Hawke; comments to author 06.10.16; revised version received 11.11.16; accepted 23.11.16; published 21.12.16.

Please cite as:
Admon L, Haefner JK, Kolenic GE, Chang T, Davis MM, Moniz MH
Recruiting Pregnant Patients for Survey Research: A Head to Head Comparison of Social Media-Based Versus Clinic-Based Approaches
J Med Internet Res 2016;18(12):e326
URL: http://www.jmir.org/2016/12/e326/
doi:10.2196/jmir.6593
PMID:28003174

Original Paper

The Role of Online Social Support in Supporting and Educating Parents of Young Children With Special Health Care Needs in the United States: A Scoping Review

Beth A DeHoff[1], MPH; Lisa K Staten[2], PhD; Rylin Christine Rodgers[3], BS; Scott C Denne[4], MD

[1]Neonatology, Department of Pediatrics, Indiana University Health Physicians, Indianapolis, IN, United States
[2]Department of Social and Behavioral Sciences, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN, United States
[3]Riley Child Development Center, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, United States
[4]Neonatology, Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Beth A DeHoff, MPH
Neonatology
Department of Pediatrics
Indiana University Health Physicians
699 Riley Hospital Dr S
RR208
Indianapolis, IN, 46202
United States
Phone: 1 3179441528
Fax: 1 3179489082
Email: badehoff@iu.edu

Abstract

Background: When parents of young children with special health care needs (CSHCN) receive their child's diagnosis, they encounter information they may not understand, emotions they may not know how to cope with, and questions about their child's immediate and long-term future that frequently lack answers. The challenge of health care providers is how to prepare parents for caring for their CSHCN, for coping with any resulting challenges, and for accessing the systems and services that can assist them.

Objective: The purpose of this work was to review evidence of the information and support needs of parents of young CSHCN and to determine whether online social support can serve as an avenue for learning and empowerment for these parents.

Methods: A scoping review identified the challenges, coping mechanisms, and support needs among parents of CSHCN, and the reach and effectiveness of digital technologies with these families and health care providers. We also conducted interviews with professionals serving parents of CSHCN.

Results: The literature review and interviews suggested that parents best learn the information they need, and cope with the emotional challenges of raising a CSHCN, with support from other parents of CSHCN, and that young parents in recent years have most often been finding this parent-to-parent support through digital media, particularly social media, consistent with the theory of online social support. Evidence also shows that social media, particularly Facebook, is used by nearly all women aged 18-29 years across racial and socioeconomic lines in the United States.

Conclusions: Parents of young CSHCN experience significant stress but gain understanding, receive support, and develop the ability to care for and be advocates for their child through parent-to-parent emotional and informational social support. Online social support is most effective with young adults of childbearing age, with social media and apps being the most useful within the theoretical framework of social support. This opens new opportunities to effectively educate and support parents of young CSHCN. Providers seeking to inform, educate, and support families of CSHCN should develop strategies to help parents find and use social support through digital resources to facilitate their emotional adjustment and practical abilities to care for and access services for their child.

(J Med Internet Res 2016;18(12):e333) doi:10.2196/jmir.6722

KEYWORDS

health communication; child; social media; health education; health resources; early childhood; disability; neonatal intensive care unit; family; maternal-child health services

Introduction

Parents of infants and young children with special health care needs (CSHCN) are often thrust into a world they know little about—a vast network of professionals, systems, and services that address the needs of children with disabilities. (In this paper, "parents" refers to guardians of all types.) At the same time, these young parents must learn to cope emotionally and practically with a child's medical and developmental needs, a situation far from the dreams of many young parents. These challenges can confront young parents when their child receives a diagnosis as a young child delayed in reaching milestones, as a newborn requiring a neonatal intensive care unit (NICU) stay for special health care needs, or even in pregnancy with a prenatal diagnosis. A key challenge for professionals who serve these families is in helping them to gain emotional support and, subsequently, the informational knowledge and skills necessary to be advocates for and achieve better health outcomes for their children. Improving social support is one way to accomplish this. Increasingly, digital communication, and social media in particular, are being used to offer online social support for parents of CSHCN. Given that research on digital communication with parents of CSHCN is lacking, we address the following question: Can we expect that social media can be an effective avenue for emotional and informational support for parents of CSHCN? In attempting to answer this question, we offer a novel exploration of how established benefits of parent-to-parent support and the theoretical framework of online social support can provide new avenues for professional partnerships to support families of CSHCN.

Challenges to Obtaining Necessary Support for Parents of CSHCN

CSHCN are defined by the US Maternal and Child Health Bureau as "those who have or are at increased risk for a chronic physical, developmental, behavioral or emotional condition and who also require health and related services of a type or amount beyond that required of children generally" [1]. When parents give birth to a CSHCN, or receive such a diagnosis prenatally or in the early years of childhood, they are immediately faced with a loss of the child they thought they would have and a tremendous amount to learn. The result is emotional stress that is overwhelming as parents sort through medical and developmental information, resources and services, varying reactions from friends and family, and even adjusting to their own new identities as parents of a CSHCN [2,3].

Shortly after a new diagnosis of a child's special health care need, parents typically receive education from their child's hospital NICU or specialist—often at one in-person appointment—about caring for their child and referrals for medical and community-level assistance. This usually occurs in the hospital just before discharge from the NICU, or in a pediatric specialty office at the end of a diagnosis appointment. At these early stages, however, parents are typically passive receivers of information as they try to process difficult information without the emotional capacity or skill sets developed to ask questions of care providers, or even to know what to ask. Mishel's theory of uncertainty in illness proposes that stressful health care experiences such as new diagnoses actually interfere with the ability of patients and families to process and understand information shared with them [4]. Only later do they start to comprehend their situation and begin to seek additional support, sometimes from care providers, but more often from family, friends, and the Internet [5]. As parents adjust to the diagnosis and start to try to make sense of how to proceed with their lives and help their child, they report that their child's medical and even social work providers have inadequate and sometimes inaccurate information about services to help the child and family at home and in their community. Initially, many parents rely on existing support networks such as family or friends who have always been their main source of all kinds of support. In many cases, these individuals lack an understanding of the child's diagnosis and resulting challenges. For emotional understanding and resources, parents have indicated that they prefer to connect with other parents of CSHCN who live in their geographic area and who have a child with the same or a similar diagnosis [6,7].

Parent-to-Parent Support as an Avenue of Social Support for Parents of CSHCN

The concepts of social support and parent-to-parent support for CSHCN both had their start in the 1970s. In 1976, Cassel concluded from both human and animal studies that social support mediates the health impact of stress [8]. In 1981, House identified four kinds of supportive actions that can help people deal with a crisis or chronic stress and allow them to cope, learn, and even grow during their challenges. This support can be emotional (being there), instrumental (doing things), informational (sharing knowledge and resources), or appraisal (helping individuals to see their stressors with more confidence in their ability to cope). Such support can provide a person with the strength to find new contacts and information in order to solve problems, the cognitive skills necessary to process that information, and the ability to better cope with stress, doubt, and fear [8]. For many young parents of CSHCN, empowerment gained from social support—particularly a combination of emotional and informational support—can lead them to successfully navigate the medical information and services necessary to help their child.

Textbox 1. Benefits of parent-to-parent support (adapted from Santelli et al [9]).

- Increased acceptance of child’s special health care needs
- Enhanced parent coping skills
- Increased self-efficacy for parents to work on problems and access services
- Rated as helpful by more than 80% of parents served
- Access to support typically unavailable from any other source
- An essential aspect of family-centered care

The idea of parent-to-parent support for families of CSHCN grew out of the movement for family-centered medical care beginning in 1971. Since that time, veteran parents of CSHCN supporting new parents has been promoted as a way for families to both reduce stress by realizing they are not alone in their struggles, and as a way to find solutions and services for their CSHCN. At first, parent-to-parent support was facilitated by health care professionals who connected a parent to another parent whose child had a similar diagnosis. Over time, organizations devoted to providing support to families of CSHCN through other parents emerged. Research into parent-to-parent support in the 1990s showed that this type of social support has several benefits, as Textbox 1 shows [9].

Parents of children with a variety of special health care needs who have been mentored by other parents of CSHCN have reported increased emotional well-being and better adaptation to their new life and identity as the parents of a CSHCN. A 2007 study of parents of CSHCN found benefits of parent-to-parent support as well, reporting that relationships with other parents helped them share experiences, feel less alone, and even find positive aspects of very painful experiences [10]. While this study focused on personal and phone support, Konrad noted that the Internet and email listservs offered additional avenues for parent-to-parent support that bridge barriers of geographic distance and rare disorders. An email support forum for parents of children with clubfoot was shown to help families, especially mothers, gain information and manage uncertainty about their child’s diagnosis [11]. Today, listservs and email are fading in popularity, while social media and mobile phone apps are becoming more popular [12]. A study of military families found a preference for online parent support via Facebook over face-to-face support, particularly among mothers [13]. Regardless of format, the reported benefits of parent-to-parent support remain consistent across time and modalities.

Traditionally, health education and social support for parents of CSHCN were available through in-person trainings and support groups. Leaders of family support nonprofit organizations for families of CSHCN responded to a common interview protocol in ways that reinforce the findings of the literature. “Support used to happen in group trainings and over kitchen tables,” said Jane Scott, assistant director of About Special Kids in Indiana, USA, a not-for-profit organization that provides parent-to-parent support for Indiana parents of CSHCN. “Now it’s social media. It’s not the same as in-person support, but it has reduced families’ isolation, and it’s especially great for rural families without options for in-person groups nearby.” Scott also noted that short message service (SMS) text messaging with families seeking information has been highly effective, with 80% of parents responding to a text message, more than twice that of parents who return a call (J Scott, oral communication, January 2016). Likewise, Jennifer Akers, project director of Family Voices Indiana, noted that parents like in-person training but, because they often lack time to attend, many interact on Facebook and “tag” her with questions. Family Voices Indiana is an Indiana nonprofit focused on parent-to-parent support for CSHCN, particularly in the areas of health care financing and advocacy. Akers said the importance of parent-to-parent organizations is that they can explain things in ways that parents understand and can apply to their lives, which she says often works better for families than information they receive from clinical and social services professionals (J Akers, written communication, January 2016). A 2014 study noted that low-income families have a particularly difficult time attending in-person support and education programs, with a recruitment rate of 31% and a high rate of attrition due to problems with child care, transportation, and time and work constraints. Swindle and colleagues found widespread access to the Internet, mobile phones, and social media among the same low-income parents, suggesting that, for low-income parents, technology may be a more accessible path to education and support than in-person programs [14]. The experiences of these nonprofit organizations serving families of CSHCN support those findings.

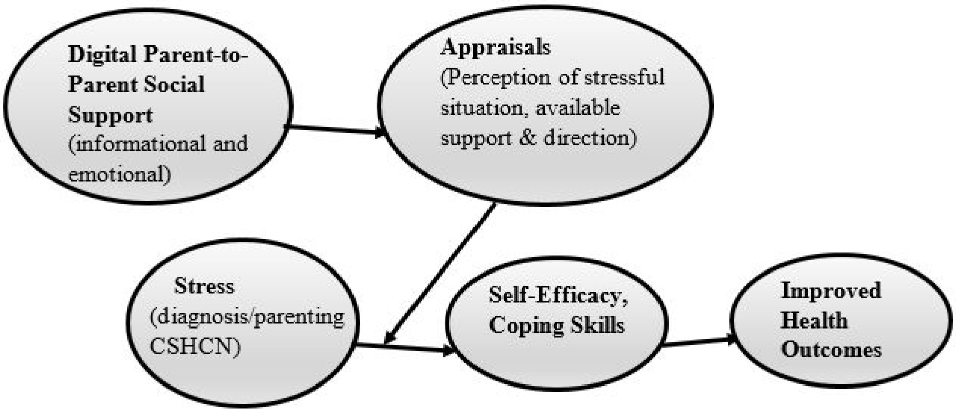
Theoretical Frameworks: Theory of Social Support

The availability of social supports helps individuals believe they are well supported, which leads them to interpret others as supportive, have a better ability to draw from past support, and have the ability to readily think about their current sources of support. This social cognitive perspective of social support [15] proposes that perceived support—an individual’s belief that she is well supported—leads to better coping skills and higher self-esteem. Individuals under stress engage in appraisals to determine whether a stressful situation is a genuine threat, and whether they have the personal and social resources to deal with the situation.

Parents of newly diagnosed young CSHCN face the stressful situation of receiving a diagnosis of their child’s condition that brings significant uncertainty and fear. However, the skills and self-confidence built through social support can empower parents to interact more positively with care providers and medical information [16]. The model shown in Figure 1 adapts Lakey and Cohen’s [15] model of the social support theory to show how emotional and informational support available from contact with other parents of CSHCN affects how parents perceive the stressful situation with their child and the supports they can count on, which mitigates stress and leads to better self-efficacy and coping skills. This belief in social support and the resulting self-efficacy ultimately allow parents to understand their child's care, seek important services for their children, and thus gain the potential to experience better health outcomes for both parent and child.

Social support can come from existing social networks (such as friends and family), new social networks (such as with other families sharing a similar diagnosis or experience), or from indigenous or community health workers. While many social service professionals focus on helping individuals strengthen their natural supports through family and friends, interventions for parents of young CSHCN also should focus on developing new social networks that exist around parenting CSHCN or dealing with specific diagnoses. New social networks often work best in response to a major life change or a specific stressor, such as a child's diagnosis. People with similar experiences can provide support based on their like journeys, providing emotional and informational support from a viewpoint only someone with similar experiences can provide [8]. Caregivers of children with life-threatening or -altering conditions share an experience unlike any that most of their friends and families can understand. Without connections to other, similar families, parents of young CSHCN often feel isolated, overwhelmed, uninformed, and uncertain. Many identify a desire to access other families with the same diagnosis for advice and support [16].

Figure 1. Digital parent-to-parent social support as a path to better outcomes for children with special health care needs (CSHCN). Adaptated from Lakey and Cohen's social support theory [15].



Informational and Emotional Support of Parents of CSHCN

Parents of children newly diagnosed with diabetes reported feeling incompetent, stressed, and anxious, and having low confidence when faced with learning how to administer home insulin injections, blood tests, and other care in a short period of time [17]. Parents indicated that the large amount of information was overwhelming and asked for repeated, short teaching sessions. In addition, parents indicated they wanted realistic, truthful information that includes not just instructions about how to care for their child but reasons why this care is needed, in ways that are practical for use in their lives at home [18]. Among parents of premature babies, parents are typically unable to actively engage in learning about their child's special health care needs immediately after diagnosis [5]. Once parents become ready to seek information, as indicated by a survey of parents of young CSHCN, their top needs for information include finding ways to help their child develop, handling the emotional impact of having a CSHCN, managing the demands on the parents' time, finding and accessing community resources, understanding their child's rights, and planning for the future [19]—needs that reflect the inextricable connection between emotional and informational social support. These needs also reflect practical, real-life coping questions that may best be answered by other parents who have learned to navigate similar needs in their own children. The 2009-2010 US National Survey of Children with Special Health Care Needs showed similar needs among parents, while also identifying additional stressors around paying for needed services, especially in the common situation of one parent leaving employment in order to care for the child [12]. Several studies have suggested that stress can be alleviated through social support, and some have concluded that learning is difficult until the stress of a situation is reduced [8,19].

The desire for social support as a means to learn about and access services is consistent with the instrumental aspect of the social support theoretical framework [15]. Given the lack of support available in one's own circle of family and friends, a lack of information in small or rural communities, and the previously referenced barriers to attending in-person training and support groups [14,20], this support is difficult to access for many families. As options for online support have grown, many parents of CSHCN are turning to online social support for information on systems and services that could benefit their child [16].

Online Social Support of Parents of CSHCN

The online social support theory [21] expands the theoretical framework of social support to include sources of support found on the Internet. This theory is based on the assertion that a change or perceived change in health can bring on acute episodes of stress that, without relief of the condition, can lead to chronic stress. The ability of a person to adapt to significant changes, such as receiving a diagnosis of a special health care need in a young child, is influenced by individual and health factors, demographics, and, increasingly, Internet use. The online social support theory proposes that online support, particularly through social media, is especially helpful for caregivers of relatives with health concerns, including parents of CSHCN [22,23].

Healthy People 2020 emphasizes health communication and information technology as key methods to improve health outcomes and increase health care equity and quality, stating specifically the role of technology in widening social support networks [24]. Digital technologies such as mobile health apps, social media groups, and other Web-based resources are 21st century solutions that directly address the need for accessible health information and communication [25]. Online support and education have been increasingly used in health care and other interventions in a variety of contexts, including with families of children with diabetes, asthma, attention-deficit/hyperactivity disorder (ADHD), and hydrocephalus [2,26,27]. The ability of these digital technologies to provide social support to parents of CSHCN achieves the traditional aims of education and information sharing through a digital avenue of online emotional and informational support. The accessibility of these technologies to parents across racial and socioeconomic lines is expanding every year.

Social media sites, particularly Facebook groups for parents of CSHCN, have high levels of intimacy and immediacy, meaning that support is available despite members' distance from one another, which naturally leads to high levels of social support. Social support is most helpful when it comes from others with similar experiences. Online peer support can provide instrumental support of information and resource sharing, as well as emotional support of caring, compassion, and inclusion [16]. Other forms of digital communication, such as blogs, listservs, chat rooms, and apps to record and share a child's milestones or health care experiences, can also provide social support but generally work best as an adjunct to Facebook groups, which are the most interactive and interpersonal formats [23]. When used by parents of CSHCN, online support has been shown to reduce isolation, increase feelings of self-control, increase parents' feelings of connection to others who understand, increase self-confidence, and lower depression and anxiety [27].

Numerous studies have concluded that mobile phone, Internet, and social media use crosses demographic lines such as income, race, and age [14,28-30]. Parents of young children have been using digital communication and social media since their teens or early adulthood, and they are already accustomed to connecting via apps, texting programs, and social media sites for social support [31]. All kinds of digital communication technologies—mobile phones, smartphones, Internet, mobile apps, Facebook, and other social media use—are most commonly used by the youngest group of adults: 18- to 29-year-olds. In this age group, the most common age group of parents of young children, shrinking social disparities in the use of these technologies are becoming even smaller. For instance, in the United States, the most prevalent users of the Internet and mobile phones are African Americans, and the most prevalent users of Facebook are Hispanics [14,28-30] (Table 1). Some studies, however, have found lower overall Internet use among minorities with low levels of education [32].

Table 1. Access to digital technologies for social support across US demographic groups[a].

Demographic groups	Access
Annual income (US $)	
High (75,000)	98% own mobile phones 78% use Facebook 84% own smartphones
Low (<30,000)	84% own mobile phones 73% use Facebook 50% own smartphones
Age range (years)	
18-29	90% use social media 86% own smartphones
30-49	83% own smartphones
All adults	90% own mobile phones 85% access Internet 64% own smartphones
Racial/ethnic/geographic groups	
African Americans	68% own smartphones
Hispanics	64% own smartphones
Whites	66% own smartphones
Rural Americans	75% view mobile health services as important

[a] Data from [14,28-30].

Provider Use of Online Social Platforms

Hospitals and health care providers are increasingly communicating digitally with one another and, less often, with patients. A survey of outpatient family practice patients revealed that 56% said they wished their doctor used social media to interact with them. Despite this interest, many clinicians have been hesitant to enter into electronic health communication with patients due to US Health Insurance Portability and Accountability Act regulatory concerns over privacy and well-publicized, inadvertent breaches of patient confidentiality by health care professionals [33]. However, most privacy breaches have been around professionals' posts on their personal social media pages, and not from official health care social media pages. In fact, health care providers are able to communicate via social media with a patient if the patient is the only one who can see it or if the patient has consented to communication via the digital platform [34]. The overriding principle in providers' involvement in digital communication efforts in general, and in social media in particular, is that the patient must be in control of their participation. It is also possible that, just as young adults are the highest users of social media, there may be generational differences in physicians' willingness to engage with patients and families online. Yet with widely available opportunities to connect online with peers beyond their own communities, and with few health care providers engaging with patients in this way, patients and families are increasingly relying less on medical providers and more on their peers for health information [35].

The findings about digital health communications in the literature are consistent with the experiences of Nerissa Bauer, MD, a pediatrician involved in research on parent engagement among children with ADHD at the Indiana University School of Medicine, Department of Pediatrics, Children's Health Services Research. Bauer has developed support groups for parents and children with ADHD and reports that getting parents to attend is a "logistical nightmare" and meets with considerable parent resistance. However, she says that parents who attend enjoy the in-person group and want to continue. Even so, she has found that parents also want information online in general and through social media particularly, echoing results from much of the literature. "Parents say they want ways to interact with health information for their kids online, but we've found that if it's not on Facebook, if it's on a different platform than what they already use, they don't use it," she said. At the same time, she said, hospitals and academic centers are extremely cautious about using Facebook for reasons relating to privacy and the expectations of parent-professional interaction online, a situation also noted in the literature. "It's definitely something to think about, and parents want it, but there are real hurdles to making it work," she said.

Currently, Bauer is working on parent-developed YouTube videos of family members talking about life with their child or grandchild with ADHD, with the idea that parents who help develop the video will share the video on their own social media channels, a video approach to parent-to-parent social support. "A mom can watch these in her pajamas at 3 AM and find families who share her experiences," Bauer said. This project is early in development and has not yet been implemented and evaluated. She also has found blogs have been a successful and controlled way to share information with parents one-way rather than interactively, but acknowledges that this strategy is better for education and information sharing than for social support (N Bauer, MD, oral communication, March 2016).

Academic literature and input from family and health leaders all point to the benefits of parent-to-parent communication and support, as well as barriers that parents of CSHCN face in finding the time and means to meet with other parents in similar situations. The aim of this scoping review was to consider what research to date has shown regarding the usefulness of digital communication, and in particular social media, in providing informational and emotional support to CSHCN.

Methods

The observations and recommendations in this paper focus on the information needs of parents of CSHCN, the barriers to face-to-face education and communication for this population, and the effectiveness of electronic media for health education and communication with parents and others using eHealth sources. It is based on a scoping review conducted in PubMed, Social Sciences Citation Index, ACM Digital Library, Education Resources Information Center (ERIC) accessed through ProQuest, and Google Scholar. Scoping reviews, like summary reviews, use rigorous and transparent research methods but focus on the findings of the reviews rather than the research used to obtain those findings [36]. We searched the databases using the following phrases, individually and in combination: “children with special needs,” “children with special health care needs,” “children with disabilities,” “parents of children with special needs,” “parents of children with special health care needs,” “parents of children with disabilities,” “parents of young children with special needs,” “parents of young children with special health care needs,” “parents of young children with disabilities,” “digital communication,” “e-health,” “electronic communication,” “electronic health education,” “health communication,” “social media”, “apps,” and “texting.” The search also included results found in PubMed’s related articles feature. We eliminated articles focusing on digital health interventions for children rather than parents, and digital health interventions for adults not targeted as parents. We also eliminated articles on health education and communication methods published prior to 2005, with searches focused on digital health education and communication articles from 2010 or later. Articles focused on the needs and challenges of parents of CSHCN were selected due to their relevancy to the topic rather than their date of publication. In each article we used, the focus was on the information and communication needs of the audience, the methods of health communication and education used, and the results for different racial and socioeconomic groups.

The collection of data also looked specifically at the use of social media, mobile phones, and other technology, and barriers to accessing various forms of education and communication, particularly among parents of CSHCN. As the literature revealed the information preferences and barriers for parents of CSHCN, it became apparent that the theories of social support and online social support were important theoretical frameworks for the subject, resulting in subsequent searches of the databases for “social support,” “online social support,” and both of these terms with the term “theory” and “theoretical framework.” Searches for health care providers’ use of social media included the terms “hospital use of social media” and “physician use of social media,” and each of these terms with “with patients” added. This was not a systematic review of the efficacy of eHealth interventions, so although the literature we present covers every subject explored in the search, it includes the sources most illustrative of these related topics rather than every article we found in the literature.

We gleaned supporting information from key-informant conversations with leaders at four Indiana organizations that serve families of CSHCN. These conversations focused on how parents engage with the organization face-to-face and online, and what strategies for informing and supporting families of CSHCN have been most successful. Another key informant was an Indianapolis pediatrician engaged in services for and research regarding children with ADHD, who discussed her work with a parent support group and her thoughts about online support for these parents.

Results

Parents of young CSHCN have enormous needs for informational and emotional support, and they often express the desire to connect with other parents for help with these needs. Juxtaposed with these needs for social support are barriers to attending in-person parent trainings or meetings, issues felt most acutely by caregivers of CSHCN who are isolated by geography or the intense care needs of their child. While research into digital communication for parents of CSHCN is limited, each study that has reported on the issue reveals that using online sources for information and support is not only possible for parents of CSHCN, it’s already happening outside of the health arena. Parents of CSHCN are already using social media, particularly Facebook, to seek out parent-to-parent online social support, and these parents report the online groups are helpful in providing informational and emotional support. These effects are consistent with the social cognitive theory of online social support [21] and social cognitive perspective of social support [15], which proposes that online support enhances participants’ belief that they are supported, affecting their perception of stressful situations and allowing them to develop coping skills that lead to better health outcomes.

In seeking effective ways to communicate with, educate, and support parents of young CSHCN, the theoretical framework of social support illustrates a way toward future efforts. For decades, parent-to-parent support has been a key pathway toward parents’ knowledge and understanding of services and supports for their CSHCN and how to access them, within the context of emotional support. This existing structure reflects the social support theory, that individuals with good social support in a crisis or time of significant change, particularly from individuals with similar experiences, are better able to cope with and move forward in the situation. Today, barriers to attending in-person programs and the availability of online social networking has moved parent-to-parent support to a largely digital venture, primarily through Facebook. This online support, consistent with the theory of online social support, can buoy parents with coping skills and self-efficacy, as well as information, to allow them to better care for and be advocates for their children.

The literature about how parents of CSHCN find support is echoed in the words of the physicians and professionals in personal interviews, and reinforced by other evidence found in randomized controlled trials about health communications. Many of these studies have concluded that texting, mobile apps, and social media are all effective ways of educating and supporting people facing serious life changes or health issues, particularly among younger adults such as parents of young children [31,37]. Following is a look at the literature involving each of these digital formats.

Mobile Phones and Texting

Texting is perhaps the first digital technology used for health communication, and it still is in use today, with uses ranging from appointment reminders, to health behavior prompts, to health information messages [38]. The ability to send and receive text messages is widespread in the United States, with 90% of all adults owning a mobile phone. Among adults 18-29 years old, 98% have a mobile phone, as do 97% of adults age 30-49 years [39]. This popularity gave rise to the text4baby program, which debuted in 2010 as a way to deliver prenatal and postpartum health messages to pregnant women and mothers who recently delivered, with messages timed to the baby's gestational age based on due date. As of February 2016, text4baby had served almost 975,000 unique pregnant and recently delivered women [40]. In 2014, a study of the use of text4baby by 943 women at an army medical center found that text messages affected the women's beliefs about health issues, particularly the importance of taking prenatal vitamins, the importance of seeing a health care provider, and the risks of drinking alcohol during pregnancy. It did not, however, seem to affect health behaviors [41]. Other studies found that text health programs had significant effects on smoking cessation and mixed effects on diabetes management and asthma management, with limited evidence and a lack of research into the long-term effects of health text programs [27,42]. Researchers have noted that studies of texting as a health intervention in general is difficult due to the variety of strategies, conditions, and subjects involved, making it unrealistic to evaluate effectiveness of texting interventions outside of small programs focused on a narrow subject, such as diabetes [38].

The literature we searched revealed no studies related to texting interventions for parents of CSHCN, and no studies relating text messaging to social support. Texting in health care has exclusively been used for one-way informational messages, which may have educational benefits but does not provide any type of emotional support, which requires some level of interaction and is, for parents of CSHCN, a necessary component of informational support [19,38].

Mobile Phones and Apps

Apps on mobile phones are some of the newest and most interactive digital modalities for education and communication. On apps, patients are not only getting health information, but also logging their weight, fitness activity, asthma peak flow numbers, insulin levels, food intake, and more, moving patients and families from passive recipients of information to active participants in their health and health care [43]. In addition, apps for mobile phones and tablets are relatively inexpensive, easily customizable to specific patient populations, and widely accessible [44]. Accessibility is growing exponentially, as 68% of Americans owned a smartphone in 2015, up from just over a third in 2011. Smartphone ownership rises with income and education. However, among the 18- to 29-year-olds, smartphone ownership is nearing saturation among all socioeconomic groups [45]. People from low-income households, as well as African American and Latino smartphone users, are more likely to be dependent on their smartphones to obtain information because many have no other access to the Internet [39]. In a study of teenagers and their caregivers in Bronx, New York, USA, most living at well below the federal poverty line and almost entirely African American and Hispanic, 85% had smartphones, and more than 70% reported accessing more than three apps a day. Singh and colleagues suggested that smartphones are becoming more affordable and accessible and, as a result, may allow public health professionals to approach elimination of health disparities in communication [46].

Each study we evaluated that used mobile apps to deliver an intervention focused on providing information and allowing users to record their own information, but lacked any aspect of social support. Any support provided involved feedback from a health care coach or other provider, providing reinforcement but lacking true emotional support. Studies of the effectiveness of apps for health education and support have been mixed. For instance, one study of an app for asthma management showed no significant impact on health outcomes [44], while another showed that app users increased peak flow, increased forced expiratory rate, and reported a high quality of life with fewer unscheduled medical visits [2]. A small study of young women aged 18-30 years in a weight loss program showed a preference for app logs over paper diaries, with 46.2% preferring apps for the program, second only to an online discussion forum. However, retention in the program (67%) was consistent with the low rates of retention and adherence seen in other weight loss programs for young women. Hutchesson and colleagues felt that the women's interest in technology interventions was promising, however, and noted that health interventions should keep up with new and changing digital platforms [47]. Another study of pregnant women compared patients who used a mobile app to record health information versus women using paper diaries. Patients using the app recorded information more frequently and were more engaged in the intervention than those using paper diaries. They reported the app was easier to use, less of a bother, and more efficient than paper journals, and app users consistently rated their satisfaction with care higher than those using paper diaries [48]. A pilot app developed for parents of high-risk infants allowed parents to record milestones and aspects of care for their baby, record certain aspects of the electronic medical record, and share what they chose with their baby's care team and on social media. This model of collaborative data collection was not tested but rather was offered as a design model. It offers promising avenues for communication and support for parents of CSHCN, as it allows parents to engage other parents with the app data on their existing digital platforms for support. However, Liu and colleagues also noted that the app's data collection feature placed a burden on parents who are already overwhelmed to collect and record data about their child, which they may find distressing, as data collection is not typically required of parents of hospitalized infants [2]. Another app for NICU families is available commercially for a fee, but it focuses on support provided to parents via a health coach rather than through peer-to-peer social support [49]. An app called NICU Companion for parents of NICU babies has been developed through the Indiana University School of Medicine [50] as in information-sharing and parent support app. Developers are expanding this platform to allow parents to track baby's progress, providing information they can then share on social media if they choose. This app is in development and could be a source for further research.

Although evidence is scant, recent studies have suggested that collaborative technology may be more readily accepted by parents than by clinicians [2]. A 2016 study found that physicians and other clinicians were resistant to using patient-collected data from an app unless it didn't take extra time and was integrated into their normal workflow. Woods and colleagues concluded that clinicians did not yet see the value of patient-collected data and may not be comfortable with the power assigned to the patient as part of his or her own care team. However, they noted that such digital collaboration could work if patients bear most of the responsibility for collecting the data and sharing it with their providers [51]. Apple has introduced its health app for iPhone, which allows individuals to track their own health, fitness, and medical information and, in some cases, share it with their doctors, fitting in neatly with emerging trends in telehealth [52]. Many apps exist or are in development to work with the Apple Health app, creating the option for shared plans of care, patient-provider communication, and other uses that could be helpful to parents of CSHCN. With very few studies focused on the use of apps by parents of CSHCN, understanding the potential and effectiveness of this technology for that population—and its potential for providing social support or enhancing collaborative family-provider partnerships—presents an opportunity for further research.

Social Media

Social media includes several forums where individuals can share thoughts and experiences with friends and followers by posting on apps and Internet sites such as Facebook, Twitter, Instagram, and other similar sites. Some studies have shown that social media use may help improve health of the users by increasing the perception of having social support and connections, and by creating patient-centered control of what is and isn't shared. Health improvements have been observed in some social media programs centered on smoking cessation and nutrition [8,16,22,23]. Potential problems with social media may arise if information provided is incorrect, or if a digital divide exists in which vulnerable populations have no access to the Internet and social media [37]. With the rapid advances of social media use and access to mobile phones, however, this digital divide appears to be shrinking. Among all US adults, 65% now use social media sites, up from 7% just 10 years ago. Nearly all (90%) of young adults aged 18-29 years use social media, and use is similar among racial and ethnic groups of all ages. People with higher incomes and higher educational levels of attainment are more likely to use social media, but these differences are far less pronounced among young adults [28]. With the explosion of social media, public health and clinical professionals are increasingly exploring its use for health communication and education [43]. Among the most receptive to social media may be young women, as social media is most commonly used by young adults and women. A 2014 study of how mothers of young children use social media revealed that young moms naturally turn to Facebook during pregnancy and early childhood to share their experiences, just as they have been doing since they were teenagers or college students. In fact, Facebook allows users to make “Expecting a Baby” a Facebook life event, and baby pictures on Facebook are so common, an app called “unbaby.me” allows users to block baby pictures from their Facebook newsfeed [31].

While there is a lack of studies regarding the use of texting and phone apps among parents of CSHCN, social media use is much more commonly studied as an intervention for this population. Social media has been shown to reduce isolation, increase emotional support, and provide useful information, addressing key challenges for parents of young CSHCN. In a 2013 study of parents of children with hydrocephalus [26], parents cited Facebook and YouTube as their preferred sources of information about hydrocephalus and their preferred way to connect with other parents. In that study, 95% of the parents indicated they used social media, and this use cut across racial and socioeconomic lines. A study of a blog and Facebook network of parents of youth with Hirschsprung disease found social media to be a successful way to connect parents of children with rare diseases with support around the world, but noted that this support lacked evidence-based health information [13]. While Facebook is the most popular social media platform, use of social media evolves quickly, requiring ongoing research and consideration of platforms such as Twitter, Instagram, and other emerging platforms of the future.

The support available on social media does more than reduce isolation. Because Facebook groups provide nearly instant access to other parents of CSHCN, parents have the opportunity to learn about systems and services available to their child and how to access them. This informal, online support group helps parents cope with and make sense of the maze of local, state, and national systems of care, as well as the overwhelming amount of information on the Internet and the network of various community services. On social media, parents help other parents navigate this complex web of information and understand what is important to do first, and eventually to be advocates for their child's needs and rights across health, education, community, and policy arenas [6]. Parents of CSHCN report they have formed meaningful relationships in social media groups, and that they feel less judged by their social media friends than by their friends and family in their “in-person” life. In social media, parents find information and support to address their most pressing needs, which often are things their family and friends don't understand. These needs include learning ways to help their children’s development, dealing with the emotions of having a CSHCN, handling time demands, finding resources in the community, planning for the future, and understanding their children’s rights. This combination of emotional and informational social support provides parents with important information and real skills that can affect outcomes for their children. In fact, support from other parents on social media has more effect on a parent’s feelings of stress than does their child’s functional level [6]. Other studies offered increasing evidence that online support, particularly through Facebook, improves feelings of well-being, which, consistent with the theory of online social support, assists parents of CSHCN through help with stress management and coping; functional support such as practical advice, information, and assistance; emotional support; and increased perceptions of being supported [16]. These benefits are routinely reported by mothers of CSHCN, as seen in one study focused on moms protesting the closing of an online support network for CSHCN, who praised the empowerment they received from the group, from information shared and emotional support [53]. Parents who receive emotional and informational support from their online friends, especially when those friends are geographically close by or whose children share similar diagnoses, eventually develop the belief they are supported and are able to summon the self-efficacy needed to advocate and pursue services for their CSHCN.

Discussion

The literature about digital platforms suggests that social media is ideally suited and already in use for supporting and informing parents of CSHCN, but that its primary use is for parent-to-parent support rather than provider-to-parent support. Texting programs have mixed success, with their most significant impact being made in health education and information sharing. The one-way nature of texting programs does not lend itself well to social support, and the multiple factors involved in texting make it difficult to assess. Because of these limitations, texting programs are not ideal for digital communication with parents of CSHCN.

Mobile phone apps are newer formats for health communication, which have shown some success with health outcomes and patient satisfaction. Most are focused strictly on providing information and collecting data, making apps an uncertain platform for social and emotional support. However, the almost total saturation of mobile phone use among young adults, indications that young adults prefer apps to paper-based health programs, barriers to in-person training and support, and the potential relevancy to interactive telehealth and telemedicine make the use of apps an intriguing possibility for parents of young CSHCN. Apps for children with serious health care needs and hospitalizations, such as babies in intensive care units, could allow parents to record important information about their child’s progress, with the potential to use the app as a tool to share updates via social media, thus creating an avenue for social support. Implementation and evaluation of an app-based intervention for parents of CSHCN would be a valuable contribution to what is currently limited literature related to this topic.

For a health care provider such as a hospital, these tools may present a way to encourage parents to connect to support from other parents via social media, without requiring medical providers to directly host a social media site for families. Finding ways to work with social media may become important for pediatric providers because social media seems to offer the most natural and effective means of online social support for parents of CSHCN.

New efforts focusing on digital communication add to the established success of parent-to-parent social support in allowing parents to share information about their child with other parents who have common experiences. Online parent-to-parent support uniquely provides the emotional and informational support that parents of CSHCN often find difficult to obtain in their own families and communities. Nonprofit organizations and, with professional precautions, larger health care organizations can develop interventions around social media strategies, while health care organizations with privacy and practical concerns about maintaining social media interventions can develop mobile phone apps that generate parent-recorded data that families can share with their online social networks, or through telemedicine connections, with their child’s doctor. Apps that allow parents to easily share their child’s progress with social networks could be a focus of future app development and research. Location-based, opt-in apps could even serve to connect parents with children who have similar diagnoses in given locations, such as in hospital cafeterias. Partnerships with community family-to-family organizations who already offer online support to families of CSHCN could be another option for health care providers, providing an avenue to reach families with information through their community partners’ social media channels. In these ways, parents learn about their children and their needs from health care professionals, share this information with their social networks, and then receive from those networks the emotional and informational support that applies to real life as a parent of a CSHCN. Exploration of potential online programs to enhance provider-parent relationships should include research into physicians’ and other medical professionals’ goals and barriers around online family support. Additionally, online support among youth with special health care needs—connecting them with other young people with similar diagnoses—would be another new area of research within the framework of online social support. In addition, efforts to further establish associations between social support and positive health outcomes can lead to greater acceptance of such programs and better options for funding.

Online social support has been shown to be effective in helping parents accept their child’s diagnosis and develop the skills they need to help their child in the days and years to come. Additional research-based evidence about online tools for communication and partnership between parents of CSHCN and physicians could advance the adoption of such tools by health care providers if such tools are shown to be effective. Digital interventions that work to inform, educate, and empower parents of CSHCN, including interventions that work within the framework of online social support, should be the focus of future research, with the goal of improving outcomes for CSHCN and their families.

Acknowledgments

Scott C Denne, MD, was funded in part by the Indiana Clinical and Translational Sciences Institute, which is supported by National Institutes of Health grant #UL1TR001108.

Conflicts of Interest

None declared.

References

1. McPherson M, Arango P, Fox H, Lauver C, McManus M, Newacheck PW, et al. A new definition of children with special health care needs. Pediatrics 1998 Jul;102(1 Pt 1):137-140. [Medline: 9714637]
2. 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]
3. Ammari T, Schoenebeck S, Morris M. Accessing social support and overcoming judgment on social media among parents of children with special needs. 2014 Jun 01 Presented at: International AAAI Conference on Web and Social Media; June 1-4, 2014; Ann Arbor, MI, USA p. 22-31 URL: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewFile/8032/8102
4. Mishel MH. Perceived uncertainty and stress in illness. Res Nurs Health 1984 Sep;7(3):163-171. [Medline: 6567948]
5. Kim UO, Basir MA. Informing and educating parents about the risks and outcomes of prematurity. Clin Perinatol 2014 Dec;41(4):979-991. [doi: 10.1016/j.clp.2014.08.015] [Medline: 25459785]
6. Ammari T, Schoenebeck S. Networked empowerment on Facebook among parents of children with special needs. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Apr 18 Presented at: CHI 2015; April 18-25, 2015; Seoul, South Korea. [doi: 10.1145/2702123.2702324]
7. Scharer K. Internet social support for parents: the state of science. J Child Adolesc Psychiatr Nurs 2005;18(1):26-35. [doi: 10.1111/j.1744-6171.2005.00007.x] [Medline: 15701096]
8. Heaney C, Israel B. Social networks and social support. In: Health Behavior and Health Education: Theory, Research and Practice. San Francisco, CA: Jossey-Bass; 2008:189-210.
9. Santelli B, Turnbull A, Higgins C. Parent to parent support and health care. Pediatr Nurs 1997;23(3):303-306. [Medline: 9220808]
10. Konrad SC. What parents of seriously ill children value: parent-to-parent connection and mentorship. OMEGA J Death Dying 2007 Oct 01;55(2):117-130. [doi: 10.2190/OM.55.2.b]
11. Oprescu F, Campo S, Lowe J, Andsager J, Morcuende JA. Online information exchanges for parents of children with a rare health condition: key findings from an online support community. J Med Internet Res 2013;15(1):e16 [FREE Full text] [doi: 10.2196/jmir.2423] [Medline: 23470259]
12. Data Resource Center for Child & Adolescent Health. 2009-2010 National Survey of Children with Special Health Care Needs. 2016 Jul 31. URL: http://www.childhealthdata.org/learn/NS-CSHCN [accessed 2016-08-01] [WebCite Cache ID 6jQWyPk2p]
13. Doty JL, Rudi JH, Pinna KL, Hanson SK, Gewirtz AH. If you build it, will they come? Patterns of internet-based and face-to-face participation in a parenting program for military families. J Med Internet Res 2016 Jun 22;18(6):e169 [FREE Full text] [doi: 10.2196/jmir.4445] [Medline: 27334833]
14. Swindle TM, Ward WL, Whiteside-Mansell L, Bokony P, Pettit D. Technology use and interest among low-income parents of young children: differences by age group and ethnicity. J Nutr Educ Behav 2014;46(6):484-490. [doi: 10.1016/j.jneb.2014.06.004] [Medline: 25087748]
15. Lakey B, Cohen S. Social support theory and measurement. In: Social Support Measurement and Intervention. New York, NY: Oxford University Press; 2000:29-52.
16. Lee R, Kvasny L. Understanding the role of social media in online health: a global perspective on online social support. First Monday 2014;19(1-6) [FREE Full text] [doi: 10.5210/fm.v19i1.4048]
17. Sullivan-Bolyai S, Bova C, Lee M, Johnson K. Development and pilot testing of a parent education intervention for type 1 diabetes: parent education through simulation-diabetes. Diabetes Educ 2012;38(1):50-57 [FREE Full text] [doi: 10.1177/0145721711432457] [Medline: 22222512]
18. Jönsson L, Hallström I, Lundqvist A. “The logic of care” - parents' perceptions of the educational process when a child is newly diagnosed with type 1 diabetes. BMC Pediatr 2012 Oct 20;12:165 [FREE Full text] [doi: 10.1186/1471-2431-12-165] [Medline: 23083125]
19. Gowen JW, Christy DS, Sparling J. Informational needs of parents of young children with special needs. J Early Intervention 1993 Jan 01;17(2):194-210. [doi: 10.1177/105381519301700209]

20. Hu C, Kung S, Rummans TA, Clark MM, Lapid MI. Reducing caregiver stress with internet-based interventions: a systematic review of open-label and randomized controlled trials. J Am Med Inform Assoc 2014 Aug 14;22:e194-e209. [doi: 10.1136/amiajnl-2014-002817] [Medline: 25125686]
21. LaCoursiere SP. A theory of online social support. ANS Adv Nurs Sci 2001 Sep;24(1):60-77. [Medline: 11554534]
22. Baum LS. Internet parent support groups for primary caregivers of a child with special health care needs. Pediatr Nurs 2004;30(5):381-8, 401. [Medline: 15587531]
23. Dietrich C. Online social support: an effective means of mediating stress. Inquiries J 2010;2(2):1-5 [FREE Full text]
24. Health communication and health information technology. Healthy People 2020. Washington, DC: U.S. Department of Health and Human Services; 2016 Dec 19. URL: https://www.healthypeople.gov/2020/topics-objectives/topic/health-communication-and-health-information-technology [accessed 2016-12-19] [WebCite Cache ID 6mso5DCMB]
25. Bender MS, Choi J, Arai S, Paul SM, Gonzalez P, Fukuoka Y. Digital technology ownership, usage, and factors predicting downloading health apps among caucasian, Filipino, Korean, and Latino Americans: the digital link to health survey. JMIR Mhealth Uhealth 2014;2(4):e43 [FREE Full text] [doi: 10.2196/mhealth.3710] [Medline: 25339246]
26. Naftel RP, Safiano NA, Falola MI, Shannon CN, Wellons JC, Johnston JM. Technology preferences among caregivers of children with hydrocephalus. J Neurosurg Pediatr 2013 Jan;11(1):26-36. [doi: 10.3171/2012.9.PEDS12208] [Medline: 23092227]
27. 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]
28. Perrin A. Social media usage: 2005-2015. Washington, DC: Pew Research Center; 2015 Oct 08. URL: http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/ [accessed 2016-08-29] [WebCite Cache ID 6k73oV2Ct]
29. Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M. Social media update 2014. Washington, DC: Pew Research Center; 2015 Jan 09. URL: http://www.pewinternet.org/2015/01/09/social-media-update-2014/ [accessed 2016-08-28] [WebCite Cache ID 6k6e06YDe]
30. Sankaranarayanan J, Sallach RE. Rural patients' access to mobile phones and willingness to receive mobile phone-based pharmacy and other health technology services: a pilot study. Telemed J E Health 2014 Feb;20(2):182-185. [doi: 10.1089/tmj.2013.0150] [Medline: 24205837]
31. Morris M. Social networking site use by mothers of young children. In: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing. 2014 Feb 15 Presented at: CSCW '14; February 15-19, 2014; Baltimore, MD p. 1272-1282 URL: http://research.microsoft.com/en-us/um/people/merrie/papers/moms_and_social_networks_cscw2014.pdf [doi: 10.1145/2531602.2531603]
32. Park E, Kim H, Steinhoff A. Health-related internet use by informal caregivers of children and adolescents: an integrative literature review. J Med Internet Res 2016;18(3):e57 [FREE Full text] [doi: 10.2196/jmir.4124] [Medline: 26940750]
33. Chretien KC, Kind T. Social media and clinical care: ethical, professional, and social implications. Circulation 2013 Apr 2;127(13):1413-1421 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.128017] [Medline: 23547180]
34. Social media compliance challenges: from HIPAA to the NLRA: social media and HIPAA privacy concerns for healthcare providers. American Health Lawyers Association, Hospitals and Health Systems Law Institute. 2013. URL: https://www.healthlawyers.org/Events/Programs/Materials/Documents/HHS13/Z_rorer.pdf [accessed 2016-12-19] [WebCite Cache ID 6k74y8gvH]
35. Sarasohn-Kahn J. The wisdom of patients: health care meets online social media. California Healthcare Foundation; 2008 URL: http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/PDF%20H/PDF%20HealthCareSocialMedia.pdf [accessed 2016-08-29] [WebCite Cache ID 6k75Eyla3]
36. Weeks LC, Strudsholm T. A scoping review of research on complementary and alternative medicine (CAM) and the mass media: looking back, moving forward. BMC Complement Altern Med 2008 Jul;8:43 [FREE Full text] [doi: 10.1186/1472-6882-8-43] [Medline: 18638413]
37. 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]
38. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
39. Smith A. U.S smartphone use in 2015. Washington, DC: Pew Research Center; 2015 Apr 01. URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2016-12-19] [WebCite Cache ID 6k76GcZPR]
40. Text4baby. Text4baby enrollment data. Arlington, VA: Voxiva Inc; 2016. URL: https://partners.text4baby.org/index.php/get-involved-pg/partners/state-organizations/2-uncategorised/105-text4baby-enrollment-data [accessed 2016-08-29] [WebCite Cache ID 6k76SLBgm]
41. Evans WD, Wallace BJ, Szekely D, Nielsen P, Murray E, Abroms L, et al. Initial outcomes from a 4-week follow-up study of the Text4baby program in the military women's population: randomized controlled trial. J Med Internet Res 2014;16(5):e131 [FREE Full text] [doi: 10.2196/jmir.3297] [Medline: 24846909]
42. Vodopivec-Jamsek V, de Jongh T, Gurol-Urganci I, Atun R, Car J. Mobile phone messaging for preventive health care. Cochrane Database Syst Rev 2012;12:CD007457. [doi: 10.1002/14651858.CD007457.pub2] [Medline: 23235643]

43. Santoro E. [Social media and medical apps: how they can change health communication, education and care]. Recenti Prog Med 2013 May;104(5):179-180. [doi: 10.1701/1291.14273] [Medline: 23748682]
44. Belisario M, 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]
45. Anderson M. Technology device ownership: 2015. Washington, DC: Pew Research Center; 2015 Oct 29. URL: http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015 [accessed 2016-08-01] [WebCite Cache ID 6jQXObzUc]
46. Singh A, Wilkinson S, Braganza S. Smartphones and pediatric apps to mobilize the medical home. J Pediatr 2014 Sep;165(3):606-610. [doi: 10.1016/j.jpeds.2014.05.037] [Medline: 24986454]
47. Hutchesson MJ, Morgan PJ, Callister R, Pranata I, Skinner G, Collins CE. Be positive be healthe: development and implementation of a targeted e-health weight loss program for young women. Telemed J E Health 2016 Jun;22(6):519-528. [doi: 10.1089/tmj.2015.0085] [Medline: 26701611]
48. Ledford CJW, Canzona MR, Cafferty LA, Hodge JA. Mobile application as a prenatal education and engagement tool: a randomized controlled pilot. Patient Educ Couns 2016 Apr;99(4):578-582. [doi: 10.1016/j.pec.2015.11.006] [Medline: 26610389]
49. NeoCare Solutions. Having a baby in the NICU isn't easy. Blue Bell, PA; 2016. URL: https://www.neocaresolutions.com/ [accessed 2016-08-29] [WebCite Cache ID 6k77jR0TJ]
50. Indiana University. NICU companion. Indianapolis, IN: The Trustees of Indiana University; 2015. URL: https://www.nicunest.org/companion/ [accessed 2016-08-29] [WebCite Cache ID 6k77rZVHl]
51. Woods SS, Evans NC, Frisbee KL. Integrating patient voices into health information for self-care and patient-clinician partnerships: Veterans Affairs design recommendations for patient-generated data applications. J Am Med Inform Assoc 2016 May;23(3):491-495. [doi: 10.1093/jamia/ocv199] [Medline: 26911810]
52. Here's what you can really do with Apple's Health app. Gear & Style CheatSheet. Asheville, NC: The Cheat Sheet; 2015 May 22 URL: http://www.cheatsheet.com/technology/the-best-new-apps-that-use-apples-healthkit.html/?a=viewall [accessed 2016-08-29] [WebCite Cache ID 6k78bPDYl]
53. Wittmeier K, Holland C, Hobbs-Murison K, Crawford E, Beauchamp C, Milne B, et al. Analysis of a parent-initiated social media campaign for Hirschsprung's disease. J Med Internet Res 2014;16(12):e288 [FREE Full text] [doi: 10.2196/jmir.3200] [Medline: 25499427]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
CSHCN: children with special health care needs
ERIC: Education Resources Information Center
NICU: neonatal intensive care unit
SMS: short message service

Edited by G Eysenbach; submitted 01.10.16; peer-reviewed by K Herron PhD, M Hanley MA; comments to author 20.10.16; revised version received 27.10.16; accepted 23.11.16; published 22.12.16.

Please cite as:
DeHoff BA, Staten LK, Rodgers RC, Denne SC
The Role of Online Social Support in Supporting and Educating Parents of Young Children With Special Health Care Needs in the United States: A Scoping Review
J Med Internet Res 2016;18(12):e333
URL: http://www.jmir.org/2016/12/e333/
doi:10.2196/jmir.6722
PMID:28007689

Viewpoint

Reimagining Human Research Protections for 21st Century Science

Cinnamon Bloss[1,2,3*], PhD; Camille Nebeker[1,3*], MS, EdD; Matthew Bietz[4], PhD; Deborah Bae[5], MPA, MBA; Barbara Bigby[6], CCRC, MA; Mary Devereaux[7], PhD; James Fowler[8,9], PhD; Ann Waldo[10], JD, CIPP; Nadir Weibel[1,11], PhD; Kevin Patrick[1,3], MS, MD; Scott Klemmer[11,12], PhD; Lori Melichar[5], PhD

[1]Center for Wireless and Population Health Systems, The Qualcomm Institute, Calit2, University of California, San Diego, La Jolla, CA, United States
[2]Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States
[3]Department of Family Medicine and Public Health, School of Medicine, University of California, San Diego, La Jolla, CA, United States
[4]Department of Informatics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, Irvine, CA, United States
[5]Robert Wood Johnson Foundation, Princeton, NJ, United States
[6]Scripps Human Research Protections Program, La Jolla, CA, United States
[7]Department of Pathology, School of Medicine, University of California, San Diego, La Jolla, CA, United States
[8]Department of Political Science, University of California, San Diego, La Jolla, CA, United States
[9]Department of Medicine, School of Medicine, University of California, San Diego, La Jolla, CA, United States
[10]Waldo Law Offices, PLLC, Washington, DC, United States
[11]Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, United States
[12]Department of Cognitive Science, University of California, San Diego, La Jolla, CA, United States
*these authors contributed equally

Corresponding Author:
Cinnamon Bloss, PhD
Center for Wireless and Population Health Systems, The Qualcomm Institute, Calit2
University of California, San Diego
9500 Gilman Drive
Atkinson Hall (mail code 0811, office 6103)
La Jolla, CA, 92093-0811
United States
Phone: 1 858 534 9595
Fax: 1 858 534 9404
Email: cbloss@ucsd.edu

Abstract

Background: Evolving research practices and new forms of research enabled by technological advances require a redesigned research oversight system that respects and protects human research participants.

Objective: Our objective was to generate creative ideas for redesigning our current human research oversight system.

Methods: A total of 11 researchers and institutional review board (IRB) professionals participated in a January 2015 design thinking workshop to develop ideas for redesigning the IRB system.

Results: Ideas in 5 major domains were generated. The areas of focus were (1) improving the consent form and process, (2) empowering researchers to protect their participants, (3) creating a system to learn from mistakes, (4) improving IRB efficiency, and (5) facilitating review of research that leverages technological advances.

Conclusions: We describe the impetus for and results of a design thinking workshop to reimagine a human research protections system that is responsive to 21st century science.

(J Med Internet Res 2016;18(12):e329) doi:10.2196/jmir.6634

KEYWORDS

ethics committees, research; biomedical research; telemedicine; informed consent; behavioral research

Introduction

Over half a century ago, in response to egregious cases of research participant mistreatment, the US government proposed prospective review of research involving human participants. This prospective review process is what we now know as the institutional review board (IRB) system. Today, IRBs are firmly entrenched within the fabric of academic research institution, with estimates putting the number of IRBs in the United States at around 6000 (I Prichard, Senior Advisor to the Director of the Office for Human Research Protections, Department of Health and Human Services; oral communication, September 2014).

IRBs have aimed to serve an important function, which is to protect human research participants. While IRBs have helped address this critical need, the IRB system has not kept pace with the evolution of research methods and practices or current and emerging trends in science and technology. The fact that the system has become antiquated calls into question whether the IRB continues to foster the protection of human research participants per the principles originally put forth in the Belmont Report [1]. New forms of research enabled by technological advances in information technology and data science appear to be particularly challenging to IRBs [2], yet clear standards to guide best practices are not well established [3-5]. We propose that the time has come to reimagine and ultimately work toward redesigning our human research protections system so that it is responsive to both the evolution of general research practices and new forms of research enabled by technological advances—what we refer to here as 21st century science. This is critical for the proper protection of research participants, ethical and efficient use of research resources, and facilitation of research insights important for human health specifically and knowledge production more generally.

A Changing Research Landscape

The IRB model was created when research was typically conducted by a single principal investigator in a single academic institution, and when data were both scarce and expensive to collect. Today, multiple principal investigator, multi-institution, and even multicountry studies are common, and such studies have resulted in unprecedented insights regarding human health. Researchers now need, or are even expected to share, data between different universities, across entities in different sectors (eg, universities, corporations, and nonprofits) and frequently across international borders. It also used to be that the scale of research was closely linked to the research methods. For instance, intervention studies were conducted with small numbers of participants in tightly controlled environments, and large-N surveys tended to collect data in ways that limited the possibility of individual identification and promoted easy anonymization. Today, expanding computational capabilities, social media, and broad research networks allow us to conduct an intervention study on Facebook with millions of participants [6], engage patients using mobile phone technology [7], study the whole genomes of thousands of individuals [8], or collect digital traces of human activity [9] at such granular levels that reidentification of individuals is possible if one possesses the right tools and expertise [10]. While traditional approaches to research require collecting only as much data as is necessary to test a hypothesis, data mining and other big data techniques derive their power from large data sets, where it may be impossible to determine, a priori, which variables will be of interest.

A Static Regulatory Environment

In contrast to the evolving research practices landscape, procedures for research oversight have been markedly static. The Common Rule, which refers to a set of regulations that specify the procedures for establishing and operating IRBs, was adopted in 1991, and the Belmont Report and the Common Rule remain the primary sources for guiding review of human research. In 2011, recognizing that these regulations had not kept pace with the evolving human research enterprise, the US Department of Health and Human Services issued an Advance Notice of Proposed Rulemaking (ANPRM) aimed at “enhancing protections for research subjects and reducing burden, delay, and ambiguity for investigators” [11]. In 2015, the ANPRM transitioned to a Notice of Proposed Rulemaking (NPRM), the next step in the process to update federal regulations [12]. The NPRM updates include, for example, new consent requirements for biological specimens, use of a central IRB for multisite studies, and changes to procedures for determining exempt versus expedited study review categories. Use of a central IRB is particularly contentious, with concerns focusing on whether protection of participants may be compromised for the increased efficiency of a single IRB [13]. Regardless, in June 2016, the US National Institutes of Health published a policy requiring single IRB review for multisite studies [14]. While the NPRM reflects important and potentially promising activity toward IRB system improvements, many have questioned or objected outright to some of the proposed changes, and even supporters have suggested they are not ideal [15]. Most recently, a report by the National Academies of Science, Engineering, and Medicine Committee on Federal Research Regulations and Reporting Requirements issued a report criticizing the NPRM, citing that the proposed changes would be detrimental to advancing research [16]. This committee recommended that the US Congress authorize the presidential appointment of a national commission to examine and update the ethical and regulatory frameworks governing human research protections. Regardless of whether and to what extent the Common Rule or principles of the Belmont Report are revised, the extent to which IRBs can keep abreast of changes in the research landscape and be responsive to studies that leverage emerging technologies remains questionable at best.

A Flawed Institutional Review Board System

There is increasing evidence that the IRB system is deeply and inherently flawed [17]. Lidz and colleagues captured the tip of the proverbial iceberg in their study of 20 IRB panels at 10 large medical institutions, where they documented 104 protocol reviews [18]. They found that IRBs consistently discussed the informed consent document, one of the Common Rule’s central mandates, and requested changes to the consent document in 88% of those cases. They also documented a disturbing observation, which was that other elements of the Common

Rule (eg, data monitoring and protection of vulnerable populations) that are intended to promote research ethics were rarely discussed. Furthermore, in studies that exceeded minimal risk of harm, 21% of reviews did not address the inclusion of strategies to minimize risk. Likewise, they noted that 50% of reviews did not compare risks and benefits, and 60% of the protocols that excluded groups of potential research subjects without explicit justification were not discussed. They also found that critical review of the research design and methodology was not uncommon, and that IRBs often requested that investigators make changes to their research design, which is typically considered outside the purview and mission of IRBs.

Taken together, these observations call into question whether IRB members are sufficiently familiar with the standards intended to guide their review of research. Furthermore, while this is the case with respect to studies that leverage traditional research methods (eg, clinical trials), these concerns are magnified when the studies under review involve emerging technologies and nontraditional methods that the IRB was not originally designed to handle and that IRB members often do not understand. Examples of such studies are those that use smartphone capabilities to measure physical activity, social media to assess adverse drug reactions, or N-of-1 genome sequencing studies for diagnosis of rare disease. Such studies raise new and nuanced ethical issues regarding participant privacy, informed consent, and data security. Some of these novel methods also inadvertently include nonparticipants [19] or “bystanders” [20] in the research record, raising potential concerns that further challenge IRB processes.

Methods

How Might We Redesign the Institutional Review Board?

In light of these issues, in January 2015 we assembled a multidisciplinary group of 11 researchers and IRB professionals drawn from academic and research institutions in San Diego, California, to consider how we might reimagine and redesign human research protections for 21st century science. The half-day workshop was set up as a brainstorming session to generate ideas for addressing IRB challenges related to review of human studies, with a particular focus on studies that leverage emerging technologies and methods. The aim was not only to stimulate creative thinking about how the existing IRB structure and process could be modified to meet the often cited challenges of the current system, but also to generate ideas for exploring entirely new ways of evaluating research to ensure that research participants are informed and protected.

Design Thinking

A central feature of this workshop was the use of design thinking strategies in the brainstorming process. Design thinking is a formal method for practical and creative resolution of problems [21] that emphasizes a phase during which the group or team focuses on generating as many ideas as possible using thoughtful prompts (eg, How might we advise as opposed to restrict? How might we simplify IRB review?). Design thinking is also considered particularly useful when the problem itself, in addition to the solution, may be unknown or ill defined at the outset of the problem-solving exercise.

Workshop Description and Stages

Workshop participants included a facilitator (SK), a cofacilitator (CB), and 9 participants (the remaining authors). A high-level goal of the session was to generate ideas for how we might reimagine and ultimately redesign the human research protections system to foster the ethical conduct of research in the changing landscape of 21st century science.

The design thinking protocol consisted of 3 primary stages. During the first stage, we asked participants to brainstorm ideas using 6 categories as prompts: (1) settings, scenarios, and steps; (2) stakeholders and extreme users; (3) utopia and dystopia; (4) change levers; (5) change agents and obstacles; and (6) things to find out. During the second stage, we asked participants to consider the ideas generated in stage 1 and to complete the sentence “How might we...?” using the stage-1 ideas as prompts. A total of 22 “How might we…?” statements were generated (see Textbox 1). From the full list of “How might we…?” statements generated in stage 2, we asked participants to select 3 ideas that they were most interested in pursuing further.

In stage 3, participants were broken into groups based on overlapping interests to further discuss and expand on specific ideas. The 5 refined “How might we…?” statements that received the most votes were (1) How might we redesign the consent form and process? (2) How might we empower researchers to protect their participants? (3) How might we learn from our efforts to protect participants? (4) How might we make the IRB system more efficient? and (5) How might we help IRBs review new forms of research enabled by technological advances? The group discussions related to each of these ideas are presented below.

Textbox 1. “How might we…?” statements.

- Start a learning health system experiment?
- Share all of our data?
- Prevent those interested in profit from taking advantage of those interested in science?
- Conduct bold experiments? (Incentivize and facilitate)
- Expedite institutional review board (IRB) review? (More appropriately classify)
- Make consent actually informed?
- Increase transparency of IRB processes and outcomes?
- Set up an appropriate surveillance system to monitor ethical violations?
- Simplify IRB review?
- Abolish IRBs?
- Reframe the IRB as a research partner rather than a research barrier?
- Increase confidence in anonymization?
- Create a learning system where IRB evolves along with research practices?
- Engage the public in research and in helping IRBs?
- Assess the true cost of the IRB system? (Direct and indirect; What are we not doing [that we should be] because of the current IRB process?)
- Collect more empirical research on the current state of the IRB?
- Create a movement around IRB?
- Influence current legislation wisely? (Start at the state level to guide national policy; eg, California Embryonic Stem Cell Research Oversight [ESCRO] committee)
- Seek an IRB waiver process?
- Include topic experts in IRB decision-making processes?
- Advise as opposed to restrict?
- Move from permission to forgiveness?

Results

Redesigning the Consent Form and Process

The ethical principle of respect for persons implies that individuals should be informed about and voluntarily consent to participate in research. How do we ensure that consent is actually informed? How do we ensure that research participants from diverse backgrounds truly understand research study risks and opportunities? In regard to the first question, one idea may be to establish mechanisms through which participants can provide real-time feedback about their experiences to researchers. These mechanisms could serve to collect empirical data regarding the clarity of consent forms and potential participants’ perceptions of risks and benefits. These data could inform and drive potential revisions to the consent form and other aspects of the research protocol. Relatedly, it is often the case that investigators write their consent forms to adhere to institutional templates, which may prompt the inclusion of content that is not relevant to or appropriate for a study. Thus, accurate and understandable descriptions of research should be encouraged in consent forms and processes, and inappropriate adherence to templates should be discouraged.

In addition, to make the informed consent process more accessible, one idea may be to think of the Creative Commons licenses [22] as a model. Similarly to the “three layers of licenses” used by Creative Commons, research studies could create three consent forms: one that contains all the legalese and scientific exposition; one in plain English that presents the facts; and a third that is simplified even further and presents risks in bullet point format. To make the process of obtaining consent culturally appropriate for underserved and underrepresented populations, community leaders, such as a *Promotor/a* in a Latino community, could be asked to help design the consent form and facilitate its use in ways that address community-specific concerns that researchers might not anticipate. Researchers could work with the community leader to help communicate these risks in a way that resonates with the community.

Empowering Researchers to Protect Participants

It may be worthwhile to consider how to construct a system of human research protections that fosters the ethical conduct of research without relying on an institution like the IRB. How might we start anew and reimagine and redesign research oversight without the traditional IRB in mind? What would an alternative system look like? One idea is to place responsibility for participant protection on the researcher rather than on the IRB. Researchers intending to engage in human-participant research could produce a document that lays out plans and risks of the research. They could then offer those documents, along with an outline of the proposed consent process, for review by their peers. Peers would be researchers in the field of relevance for the research. These documents could be posted on the Web in the same way clinical trials are registered; not to get approval but to create a public record of the research. Peers who review the documents might be accredited with some type of certification in human research protections, although an open question would be what entity would design and provide such certification (and how such an entity would look different from a traditional IRB). Obtaining this certification and participating in this process could be incentivized for researchers by considering these activities to be professional service required for career advancement and academic promotion. In this scenario, responsibility for ethical conduct during the study would be shared by both the researchers and the peers who agreed that the plan would adequately protect participants.

To make it easier to create high-quality plans, researchers could consult a Web-based resource similar to Stack Overflow [23], a resource that software developers often use to obtain quick answers from experts about specific technical issues. With this resource, the median response time is 11 minutes [24], and the people responding are rated, which provides information pertaining to their credibility and expertise. Using this Web-based resource, within a few hours, researchers posing questions such as “How do I ensure that I won’t cause harm by asking this interview or survey question?” would receive answers from researchers who have been rated in terms of experience and expertise in human research protections. Elements of the plans could ultimately become like “protection modules” that could be swapped in and out of consent forms and research protocols, drawing attention to highly ranked modules. We note, however, that such a solution would require an active community with a critical mass of users, which may or may not be realistic depending on whether the IRB process ever became truly standardized. Importantly, if such a system were found to be feasible, it is an approach that could be coupled with a system that punishes offenders (see below).

Reinforcement and Learning From Experience

This notion also begins with the premise that the burden to protect participants be shifted to the researcher rather than remain with an IRB or other regulatory body. How might we simultaneously reduce the bureaucratic burden of IRBs for researchers, particularly those conducting low-risk studies, and, at the same time, improve protection for research participants? In addition, how might we transform universities into learning ethics institutions that continuously improve their capacity to conduct ethical research [25]? One model for doing this could be the US Federal Aviation Administration’s Aviation Safety Reporting System [26]. Pilots who have a “bad” landing or make another safety-related error who self-report their mistake are spared from punishment, but those who do not report it themselves are penalized if someone elects to report [27]. Analogously, as an alternative to an IRB, in this system, researchers who create a protocol they believe to be safe, who then observe a harm during the research and who report that harm to their university or institution, present an opportunity for the research institution and community to learn how to prevent future harm. This expectation would be reinforced because, if the harm were to be reported by anyone else, including research staff or the research participant, the researcher would be sanctioned. That being said, there are clearly potential risks of supplanting the traditional IRB with a system entirely driven by researcher self-regulation. There could be conflicts between the researchers’ mandate to conduct studies and publish them and their mandate to protect participants, thus creating the opportunity for bias, the perception of bias, or, in extreme cases, maleficence. A system of researcher self-regulation would need to carefully consider and guard against these potential threats.

Increasing Efficiency of the Institutional Review Board

We suggest that, in order to improve the IRB process, it is essential to understand its costs, both direct and indirect. How might we collect and analyze empirical data on costs of the IRB system? Obvious tangible costs associated with the IRB system include salaries of personnel, IRB fees, space and infrastructure costs, and fees paid for training, education, and accreditation. In addition, for researchers, costs include the amount of time required for study staff to prepare and process a study protocol through the IRB. Depending on the institution and the type of protocol, IRB submissions can be extremely time intensive to prepare, which is an opportunity cost in terms of other ways in which that time could be spent. For research participants, costs include the time and cognitive effort needed to understand increasingly complex and bureaucratic consent forms. There are also less-tangible costs related to the broader public health caused by unnecessary delays to research imposed by IRBs.

One idea to increase efficiency may be to use the “Cooperative Research” process (see 45 CFR 46.114 [28]) to reduce the multiple IRB review of multisite studies and to use the “exempt” category to a greater degree, as it was intended. The exempt category is frequently appropriate for the vast majority of social and behavioral science studies, yet it is underused, which leads to delays in review and approval [29] and, thus, wasted resources. In addition, IRBs could take care to ensure that the process of review for exempt studies is reasonable and truly reflects their low-risk nature. Interestingly, exempt research, according to US federal regulations, does not need to be verified or reviewed by IRB staff. If institutions permitted, determining exempt status could be made the responsibility of the researcher. Overall, the idea that the bureaucracy of the IRB creates a significant burden to the research enterprise while producing unclear or intangible benefits to research participants is consistent with the purported rationale cited for the development of the proposed revisions to the Common Rule in the form of the NPRM [11] discussed above. We suggest that the IRB may benefit from an analysis of costs and benefits of its own activities, much like it does with the studies it oversees.

Review of Research That Leverages Technological Advances

New forms of research enabled by technological advances in information technology, data science, and other fields appear to be particularly challenging to IRBs. How might we develop resources that would facilitate appropriate review of 21st century science? The California Institute for Regenerative Medicine (CIRM) research oversight process could serve as a model. In CIRM 1.0, a committee separate from the IRB called the Embryonic Stem Cell Research Oversight (ESCRO) committee was formed to review stem cell research. Recognizing that few IRB members would have sufficient expertise to provide a meaningful review, the ESCRO committee, which is composed of scientists and a community representative, serves in an advisory capacity to the IRB. Such a model could be replicated for studies using emerging technologies about which IRBs may be similarly unfamiliar to ensure that experts are involved in the review.

For example, mobile, visual imaging, pervasive sensing, and geolocation tracking technologies present new ethical and regulatory challenges [20]. For instance, visual imaging using wearable sensors have made it possible for researchers to measure physical activity, diet, travel, and the settings in which these behaviors occur using a first-person point-of-view wearable camera. Given the increasing interest in these methods for studying behavior “in the wild,” we anticipate increased research using visual methods, which raises privacy concerns and issues related to the rights of bystanders. Likewise, with wearable sensors, mobile phone transmission, and analytics in the cloud, health information can be captured continuously in real time. Location tracking technologies provide spatial data and the opportunity for assessing the context in which behavior is occurring, as well as identifying underlying spatial relationships such as clustering or transmission pathways. These data are fine grained and specific down to the exact longitude and latitude at a given point in time. Standards for how these data are transmitted, stored, and shared are necessary, since the introduction of the US Health Insurance Portability and Accountability Act, in most cases (at least at present), does not apply. A virtual network composed of researchers, technologists, and bioinformatics experts may prove to be a workable solution to augment or replace the traditional IRB review process resulting in an informed and meaningful human protections review of 21st century science.

Discussion

In this paper we imagine, and offer some ideas for the design of, a progressive, responsive, and nimble human research protections system. By encouraging broad and innovative ideas, the design thinking method not only opens up new avenues for exploration, but also provides clarity about some of the shortcomings of our current IRB system. The workshop described here aimed to stimulate creative thinking about how the existing IRB structure could be improved, while also generating ideas for entirely new ways of protecting research participants. Clearly, some of the ideas presented here are more feasible than others. For example, it may be more realistic to encourage IRBs to exploit the current regulations and use the “exempt” category more appropriately and frequently. Alternatively, imagining a completely different review process that would replace the traditional IRB entirely may be less acceptable and would likely create new problems. If the proposed NPRM is adopted, we note that some research may be excluded from the traditional IRB review yet may benefit from an “ethics consultation” process to avoid making mistakes that an IRB may have detected. It would be valuable to estimate the cost savings realized by implementing the new practices that reduce the burden on both IRBs and researchers without compromising human research protections. Clearly, further debate by stakeholders is necessary to develop these and other ideas into concrete recommendations to advance applied human research ethics.

Sometimes, systems are so entrenched in ways of doing things that change from within is not possible and disruptive external approaches are required (eg, Uber as an alternative to transportation via taxis, or specialty charter schools as an alternative to traditional public schools). The IRB may be a system in need of disruption. Using the design thinking method fostered the development of “outside the box” ideas that may improve research participant protections and the IRB structure. As such, we initiated the exercise described in this paper and seek to share the process and results with the greater research community. The IRB system will need to be updated or possibly even reinvented in order to be responsive to technological advances of recent decades that have enabled new forms of research. These advances have created challenges to our current system that the NPRM will not likely solve. Pilot research programs that test-drive the reform ideas presented here, or perhaps other ideas, would be worthwhile and informative as the research community considers how to intervene and make healthy what many believe is an ailing human research protections system.

Acknowledgments

We thank the Robert Wood Johnson Foundation for providing support for the workshop that was the source of the ideas presented. We also thank The Qualcomm Institute, the University of California, San Diego Division of the California Institute for Telecommunications and Information Technology (Calit2), for providing the space in which to hold the meeting. We also thank Ms Katrina Ansert for taking detailed notes during the workshop, Mr Johnny Nguyen for administrative support, and Ms Cynthia Cheung for assistance with manuscript preparation.

This work was supported by two grants from the Robert Wood Johnson Foundation (“The Health Data Exploration [HDE] project,” principal investigator: KP, #71693, 2013-2017, and “Connected and Open Research Ethics [CORE],” principal investigator: CN, #72876, 2015-2017) and a grant from the National Human Genome Research Institute (“Impact of Privacy Environments for Personal Health Data on Patients”, principal investigator: CB, R01 HG008753, 2015-2018).

Conflicts of Interest

None declared.

References

1. Zucker B. The Belmont Report. In: Wiley Encyclopedia of Clinical Trials. New York, NY: John Wiley & Sons; Dec 14, 2007.
2. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med 2016 Dec;6(4):577-586. [doi: 10.1007/s13142-016-0426-4] [Medline: 27688250]
3. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. JMIR Mhealth Uhealth 2015;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
4. Eysenbach G. Towards ethical guidelines for e-health: JMIR theme issue on eHealth ethics. J Med Internet Res 2000;2(1):E7 [FREE Full text] [doi: 10.2196/jmir.2.1.e7] [Medline: 11720926]
5. Conway M. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. J Med Internet Res 2014;16(12):e290 [FREE Full text] [doi: 10.2196/jmir.3617] [Medline: 25533619]
6. Bond RM, Fariss CJ, Jones JJ, Kramer AD, Marlow C, Settle JE, et al. A 61-million-person experiment in social influence and political mobilization. Nature 2012 Sep 13;489(7415):295-298 [FREE Full text] [doi: 10.1038/nature11421] [Medline: 22972300]
7. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for schizophrenia: patient engagement with a mobile phone intervention following hospital discharge. JMIR Ment Health 2016;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
8. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012 Nov 1;491(7422):56-65 [FREE Full text] [doi: 10.1038/nature11632] [Medline: 23128226]
9. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May 12;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
10. de Montjoye MY, Radaelli L, Singh VK, Pentland AS. Identity and privacy. Unique in the shopping mall: on the reidentifiability of credit card metadata. Science 2015 Jan 30;347(6221):536-539. [doi: 10.1126/science.1256297] [Medline: 25635097]
11. [No authors listed]. Human subjects research protections: enhancing protections for research subjects and reducing burden, delay, and ambiguity for investigators. Federal Register 2011 Jul 26;76(143):44512-44531 [FREE Full text]
12. [no authors listed]. Notice of proposed rulemaking: federal policy for the protection of human subjects. Federal Register 2015.
13. Klitzman R. How local IRBs view central IRBs in the US. BMC Med Ethics Jun 23 2011:12. [doi: 10.1186/1472-6939-12-13] [Medline: 21699725]
14. National Institutes of Health. Final NIH Policy on the use of a single institutional review board for multi-site research. 2016. URL: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-094.html) [accessed 2016-10-31] [WebCite Cache ID 6ltef8okD]
15. Emanuel E. Reform of clinical research regulations, finally. N Engl J Med 2015 Oct 10;373(24):2296-2299. [doi: 10.1056/NEJMp1512463] [Medline: 26536071]
16. Committee on Federal Research Regulations and Reporting Requirements: A New Framework for Research Universities in the 21st Century, Committee on Science, Technology, and Law, Board on Higher Education and Workforce, Policy and Global Affairs, National Academies of Sciences, Engineering, and Medicine. Optimizing the Nation’s Investment in Academic Research: A New Regulatory Framework for the 21st Century. Washington, DC: National Academies Press; Jul 2016.
17. Schneider C. The Censor's Hand. Cambridge, MA: MIT Press; 2015.
18. Lidz C, Garverich S. What the ANPRM missed: additional needs for IRB reform. J Law Med Ethics Summer 2013;41(2):390-396. [doi: 10.1111/jlme.12050] [Medline: 23802892]
19. Bloss C. Does family always matter? Public genomes and their effect on relatives. Genome Med 2013;5(12):107. [doi: 10.1186/gm511] [Medline: 24342550]
20. Nebeker C, Linares-Orozco R, Crist K. A multi-case study of research using mobile imaging, sensing and tracking technologies to objectively measure behaviorthical issues and insights to guide responsible research practice. J Res Admin 2015;46(1):118-137.
21. Brown T. Change by Design: How Design Thinking Transforms Organizations and Inspires Innovation. New York, NY: HarperCollins; 2009.
22. Creative Commons. About the licenses: what our licenses do. Mountain View, CA: Creative Commons URL: https://creativecommons.org/licenses/ [accessed 2016-10-19] [WebCite Cache ID 6lNd6mMfm]
23. Stack Overflow. New York, NY: Stack Exchange Inc; 2016. URL: http://stackoverflow.com/ [accessed 2016-12-13] [WebCite Cache ID 6mjRnc3eP]

24. Mamykina L, Manoim B, Mittal M, Hripcsak G, Hartmann B. Design lessons from the fastest q&a site in the west. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI '11; May 7-12, 2011; Vancouver, BC, Canada p. 2857-2866. [doi: 10.1145/1978942.1979366]
25. Senge P. The Fifth Discipline: The Art and Practice of the Learning Organization. New York, NY: Broadway Business; 2006.
26. NASA. Aviation Safety Reporting System. URL: https://asrs.arc.nasa.gov/ [accessed 2016-10-19] [WebCite Cache ID 6lNcRUWUZ]
27. Norman D. The Design of Everyday Things. New York, NY: Doubleday; 1988.
28. Office for Human Research Protections. Code of Federal Regulations, Title 45: Public Welfare, Department of Health and Human Services. Part 46: protection of human subjects. Rockville, MD: Office for Human Research Protections; 2010 Jan 15. URL: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html# [accessed 2016-12-13] [WebCite Cache ID 6mjSREPpp]
29. Schrag Z. Ethical Imperialism: Institutional Review Boards and the Social Sciences, 1965-2009. Baltimore, MD: Johns Hopkins University Press; 2010.

Abbreviations

ANPRM: Advance Notice of Proposed Rulemaking
CIRM: California Institute for Regenerative Medicine
ESCRO: Embryonic Stem Cell Research Oversight
IRB: institutional review board
NPRM: Notice of Proposed Rulemaking

Edited by G Eysenbach; submitted 08.09.16; peer-reviewed by J Hawkins, D Maslove; comments to author 19.10.16; revised version received 09.11.16; accepted 23.11.16; published 22.12.16.

Please cite as:
Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, Fowler J, Waldo A, Weibel N, Patrick K, Klemmer S, Melichar L
Reimagining Human Research Protections for 21st Century Science
J Med Internet Res 2016;18(12):e329
URL: http://www.jmir.org/2016/12/e329/
doi:10.2196/jmir.6634
PMID:28007687

Original Paper

Association Between Physician Online Rating and Quality of Care

Kanu Okike[1*], MPH, MD; Taylor K Peter-Bibb[2*]; Kristal C Xie[3*]; Okike N Okike[4*], MD

[1]Department of Orthopedic Surgery, Kaiser Permanente Moanalua Medical Center, Honolulu, HI, United States
[2]University of Colorado, Boulder, CO, United States
[3]Iolani School, Honolulu, HI, United States
[4]Department of Patient Experience, University of Massachusetts Memorial Healthcare, Worcester, MA, United States
[*]all authors contributed equally

Corresponding Author:
Kanu Okike, MPH, MD
Department of Orthopedic Surgery
Kaiser Permanente Moanalua Medical Center
3288 Moanalua Road
Honolulu, HI, 96821
United States
Phone: 1 808 432 7326
Fax: 1 808 432 8330
Email: okike@post.harvard.edu

Abstract

Background: Patients are increasingly using physician review websites to find “a good doctor.” However, to our knowledge, no prior study has examined the relationship between online rating and an accepted measure of quality.

Objective: The purpose of this study was to assess the association between online physician rating and an accepted measure of quality: 30-day risk-adjusted mortality rate following coronary artery bypass graft (CABG) surgery.

Methods: In the US states of California, Massachusetts, New Jersey, New York, and Pennsylvania—which together account for over one-quarter of the US population—risk-adjusted mortality rates are publicly reported for all cardiac surgeons. From these reports, we recorded the 30-day mortality rate following isolated CABG surgery for each surgeon practicing in these 5 states. For each surgeon listed in the state reports, we then conducted Internet-based searches to determine his or her online rating(s). We then assessed the relationship between physician online rating and risk-adjusted mortality rate.

Results: Of the 614 surgeons listed in the state reports, we found 96.1% (590/614) to be rated online. The average online rating was 4.4 out of 5, and 78.7% (483/614) of the online ratings were 4 or higher. The median number of reviews used to formulate each rating was 4 (range 1-89), and 32.70% (503/1538) of the ratings were based on 2 or fewer reviews. Overall, there was no correlation between surgeon online rating and risk-adjusted mortality rate (P=.13). Risk-adjusted mortality rates were similar for surgeons across categories of average online rating (P>.05), and surgeon average online rating was similar across quartiles of surgeon risk-adjusted mortality rate (P>.05).

Conclusions: In this study of cardiac surgeons practicing in the 5 US states that publicly report outcomes, we found no correlation between online rating and risk-adjusted mortality rates. Patients using online rating websites to guide their choice of physician should recognize that these ratings may not reflect actual quality of care as defined by accepted metrics.

(J Med Internet Res 2016;18(12):e324) doi:10.2196/jmir.6612

KEYWORDS

online reviews; cardiac surgery; physician quality

Introduction

Consumers have long used reviews of goods and services to inform their choices. Recently, these trends have spread to the health care arena in the form of online physician review websites [1-14]. According to a recent survey, 65% of respondents were aware of physician rating websites, and 35% had sought online physician reviews within the past year [15]. The survey also found that these online reviews were influential: among those who sought physician ratings information online, 35% reported selecting a physician based on good ratings and 37% reported avoiding a physician with bad ratings [15].

While patients are increasingly using physician review websites to find “a good doctor,” it remains unclear whether online physician ratings actually reflect quality of care. Segal et al analyzed online ratings in relation to surgeon case volume, which they considered to be a proxy for quality of care, and found no correlation between numerical rating and number of procedures performed [13]. Similarly, Gao and colleagues analyzed ratings from the RateMDs.com website in comparison with data obtained from the Virginia Board of Medicine, and found no correlation between physician rating and malpractice claims [8]. However, to our knowledge, no prior study has examined the relationship between online ratings and an accepted measure of quality.

The purpose of this study was to assess the degree to which online physician ratings reflect quality of care. In the US states of New York, New Jersey, Massachusetts, Pennsylvania, and California—which together account for over one-quarter of the US population [16]—risk-adjusted mortality rates are publicly reported for all cardiac surgeons. By analyzing the online ratings of these surgeons in comparison with their clinical outcomes, we sought to assess the degree to which online ratings correlate with quality of care.

Methods

In June 2015, we accessed the cardiac surgeon “report cards” for all 5 states that publicly report risk-adjusted cardiac surgery mortality rates (ie, California [17], Massachusetts [18], New Jersey [19], New York [20], and Pennsylvania [21]). From the online reports, we recorded the names of all cardiac surgeons practicing in these states, as well as their institutions. For each surgeon listed, we also recorded the 30-day risk-adjusted mortality rate following isolated coronary artery bypass graft (CABG) surgery.

To calculate the risk-adjusted mortality rate, the observed mortality rate is divided by the expected mortality rate and then multiplied by the statewide mortality rate. (For reference, the observed mortality rate is the observed number of deaths divided by the total number of cases, and the expected mortality rate is the sum of predicted probabilities of death for all patients divided by the total number of patients.)

For each surgeon listed in the state reports, we conducted Internet-based searches between July and September 2015 to determine his or her online rating(s). Searches were conducted using surgeon name, location, and specialty. For each online rating identified, we recorded the name of the website, the overall rating, and the number of reviews used to formulate the rating. Online ratings were out of 5. The individuals performing these searches (TKPB and KCX) were blinded to the surgeons’ clinical outcomes.

We assessed the association between surgeon online rating and risk-adjusted mortality rate using the Pearson correlation coefficient. In addition, surgeons were grouped on the basis of average online rating, and risk-adjusted mortality rates were compared using the Student *t* test. Surgeons were also grouped on the basis of risk-adjusted mortality rate quartile, and online ratings were compared using Student *t* test. *P*<.05 was considered statistically significant, and all tests were 2-sided. Statistical analysis was performed using SAS version 9 (SAS Institute Inc).

Results

There were 614 cardiac surgeons with risk-adjusted mortality rates listed in the 5 state reports (California: 271; Massachusetts: 52; New Jersey: 36; New York: 135; and Pennsylvania: 120). For all states combined, the average 30-day risk-adjusted mortality rate after isolated CABG surgery was 1.68% (SD 1.98, median 1.22%, range 0.00%-16.98%).

We found 96.1% (590/614) of the surgeons to be rated online, including from Healthgrades (n=540) [22], Vitals (n=495) [23], UCompareHealthCare (n=366) [24], and RateMDs (n=103) [25]. We found that 74 of the surgeons were rated on a single website, while 170 were rated on 2 websites, 266 were rated on 3 websites, and 80 were rated on 4 or more websites. The average online rating for the cardiac surgeons was 4.4 on a scale of 1-5, with 1 being the lowest score and 5 being the highest score obtainable. As Table 1 shows, 78.7% (483/614) of the scores were 4 out of 5 or better. The median number of reviews per surgeon was 4, with a wide range (1-89 reviews).

Figure 1 depicts a scatterplot of surgeon risk-adjusted mortality rate versus average online rating. Surgeon online rating did not correlate with risk-adjusted mortality rate (Pearson correlation coefficient –.06, *P*=.13). Risk-adjusted mortality rates were similar for surgeons across categories of average online rating (*P*>.05; Figure 2). Similarly, surgeon average online rating was similar across quartiles of surgeon risk-adjusted mortality rate (*P*>.05; Table 2).

Table 1. Average online ratings[a] of cardiac surgeons who had risk-adjusted mortality rates listed, July-September 2015.

Average online rating	n (%)
5.00	159 (25.9)
4.00-4.99	324 (52.8)
3.00-3.99	94 (15.3)
2.00-2.99	8 (1.3)
1.00-1.99	5 (0.8)
Not rated online	24 (3.9)
Total	614 (100.0)

[a] Ratings are out of 5.

Figure 1. Average online rating versus risk-adjusted mortality rate. Ratings are out of 5.

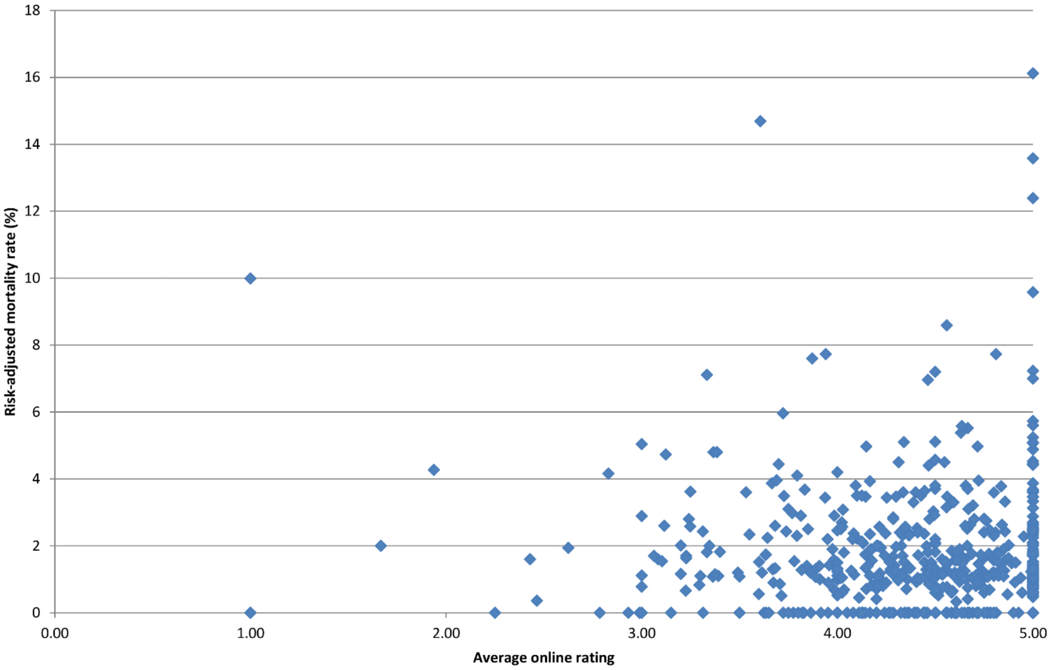


Figure 2. Risk-adjusted mortality rate, by average online rating. Note that the categories of average online rating differ in size. Error bars indicate 95% CIs, which vary in magnitude due to the number of ratings in each category (n=13 for 1.00-2.99, n=94 for 3.00-3.99, n=324 for 4.00-4.99, and n=159 for 5.00). Ratings are out of 5. There were no significant differences between the groups (*P*>.05).

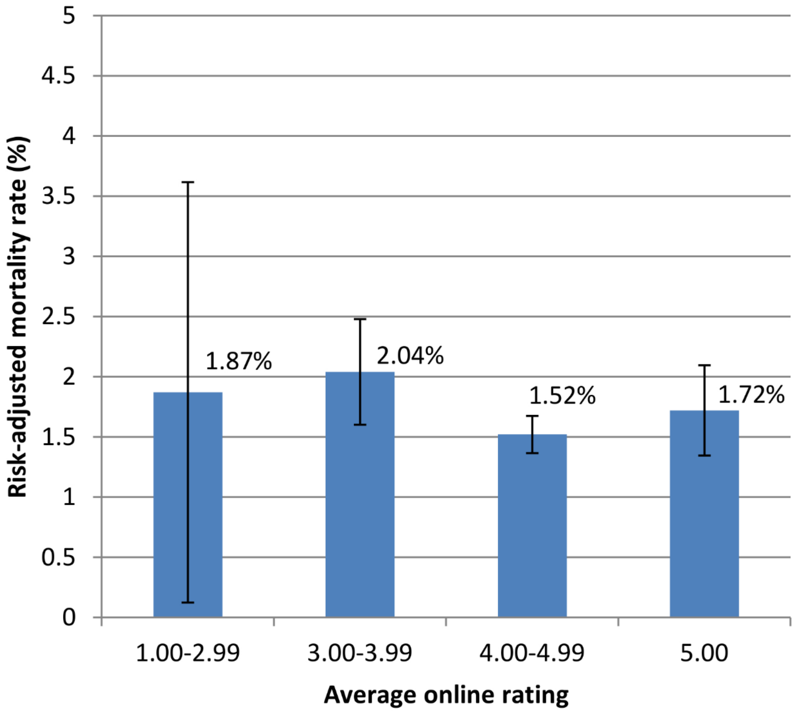


Table 2. Average online rating, by quartile of surgeon risk-adjusted mortality rate.

Surgeon risk-adjusted mortality rate			Surgeon average online rating[a]		
Quartile	Range	n	Mean	SD	*P* value[b]
Very low	0.00%-0.41%	148	4.4	0.7	(reference)
Low	0.45%-1.20%	147	4.5	0.5	.34
Medium	1.23%-2.31%	148	4.4	0.6	.49
High	≥2.34%	147	4.4	0.7	.40

[a] Ratings are out of 5.

[b] Compared with surgeons with risk-adjusted mortality rates categorized as “very low.”

Discussion

In this study of cardiac surgeons practicing in the 5 US states that publicly report outcomes, we found no correlation between online rating and risk-adjusted mortality rates.

We are not aware of any prior study assessing the correlation between physician online rating and accepted measures of quality. However, 2 prior studies have examined the relationship between patients’ subjective assessments of care and objective measures of quality. In these 2 studies, both of which were conducted among individuals over the age of 65 years, the subjective ratings given by patients were not found to correlate with the accepted quality measures [26,27].

Our study is not without its limitations. We used 30-day risk-adjusted mortality rates to measure quality, and it is possible that our results could have differed had we examined long-term mortality rates or rates of major morbidity (such as renal failure or stroke). However, 30-day risk-adjusted mortality is the most commonly accepted measure of quality in the field [28]. In addition, since we investigated cardiac surgeons in 5 US states, it is unclear whether the findings can be generalized to other fields of medicine or other locations.

For physicians, who have long argued that online ratings do not reflect clinical competency [29], the results of our study may not be surprising. However, our findings serve as a reminder that the provision of high-quality medical care may not necessarily translate into higher online ratings.

Our study also has important implications for patients. Consumers are increasingly using online reviews to guide their selection of goods and services, and health care is no exception [15]. Based on the results of our study, patients using online rating websites to guide their choice of physician should recognize that these ratings may not reflect actual quality of care as defined by accepted metrics. In contrast, they may be more reflective of factors such as clinic wait times [30] or bedside manner [31].

In summary, this study of cardiac surgeons practicing in the 5 US states that publicly report outcomes found no correlation between online rating and risk-adjusted mortality rates. Patients using online rating websites to guide their choice of physician should recognize that these ratings may not reflect actual quality of care as defined by accepted metrics.

Acknowledgments

The authors gratefully acknowledge Mr Donovan Delgado for his assistance with data collection. Publication of this paper was funded by the University of Colorado Boulder Libraries Open Access Fund.

Conflicts of Interest

K Okike receives educational meeting support from DePuy, Stryker Corporation, Synthes, and Zimmer, and teaching honoraria from Synthes. TK Peter-Bibb, KC Xie, and ON Okike report no conflicts of interest.

References

1. Atkinson S. Current status of online rating of Australian doctors. Aust J Prim Health 2014;20(3):222-223. [doi: 10.1071/PY14056] [Medline: 24852125]
2. Bakhsh W, Mesfin A. Online ratings of orthopedic surgeons: analysis of 2185 reviews. Am J Orthop (Belle Mead NJ) 2014 Aug;43(8):359-363. [Medline: 25136868]
3. Black EW, Thompson LA, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. Inform Prim Care 2009;17(4):249-253 [FREE Full text] [Medline: 20359403]
4. Detz A, López A, Sarkar U. Long-term doctor-patient relationships: patient perspective from online reviews. J Med Internet Res 2013;15(7):e131 [FREE Full text] [doi: 10.2196/jmir.2552] [Medline: 23819959]
5. Ellimoottil C, Hart A, Greco K, Quek ML, Farooq A. Online reviews of 500 urologists. J Urol 2013 Jun;189(6):2269-2273. [doi: 10.1016/j.juro.2012.12.013] [Medline: 23228385]
6. 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]
7. Emmert M, Meier F, Heider A, Dürr C, Sander U. What do patients say about their physicians? an analysis of 3000 narrative comments posted on a German physician rating website. Health Policy 2014 Oct;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]
8. 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]
9. 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 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
10. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012 Jun;27(6):685-692 [FREE Full text] [doi: 10.1007/s11606-011-1958-4] [Medline: 22215270]
11. Merrell JG, Levy BH, Johnson DA. Patient assessments and online ratings of quality care: a “wake-up call” for providers. Am J Gastroenterol 2013 Nov;108(11):1676-1685. [doi: 10.1038/ajg.2013.112] [Medline: 24192941]
12. Sabin JE. Physician-rating websites. Virtual Mentor 2013 Nov;15(11):932-936 [FREE Full text] [doi: 10.1001/virtualmentor.2013.15.11.ecas2-1311] [Medline: 24257083]
13. Segal J, Sacopulos M, Sheets V, Thurston I, Brooks K, Puccia R. Online doctor reviews: do they track surgeon volume, a proxy for quality of care? J Med Internet Res 2012;14(2):e50 [FREE Full text] [doi: 10.2196/jmir.2005] [Medline: 22491423]
14. Wallace BC, Paul MJ, Sarkar U, Trikalinos TA, 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]
15. 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]
16. United States Census Bureau. Table 1: annual estimates of the resident population for the United States, regions, states, and Puerto Rico: April 1, 2010 to July 1, 2015. 2015. URL: http://www.census.gov/popest/data/state/totals/2015/tables/NST-EST2015-01.xlsx [accessed 2016-12-11] [WebCite Cache ID 6mh8BL52g]
17. State of California Office of Statewide Health Planning and Development. CABG surgery in California. 2016. URL: http://www.oshpd.ca.gov/HID/CABG-Report.html [accessed 2016-08-21] [WebCite Cache ID 6kGAO8Yfr]
18. Massachusetts Data Analysis Center (Mass-DAC). Cardiac study: annual reports. 2016. URL: http://www.massdac.org/index.php/reports/cardiac-study-annual/ [accessed 2016-08-22] [WebCite Cache ID 6kGAbeP8t]
19. State of New Jersey Department of Health. New Jersey hospital performance report. URL: http://web.doh.state.nj.us/apps2/hpr/index.aspx [accessed 2016-08-22] [WebCite Cache ID 6kGAjcxr5]

20. New York State Department of Health. Cardiovascular disease data and statistics. URL: https://www.health.ny.gov/statistics/diseases/cardiovascular/ [accessed 2016-08-22] [WebCite Cache ID 6kGAoBTI1]
21. Pennsylvania Health Care Cost Containment Council. Public reports: cardiac care. URL: http://www.phc4.org/reports/cabg/ [accessed 2016-08-22] [WebCite Cache ID 6kGAxMldA]
22. Healthgrades: better health begins here. Denver, CO: Healthgrades Operating Company; 2016. URL: https://www.healthgrades.com [accessed 2016-12-09] [WebCite Cache ID 6mh8ybiUL]
23. Vitals: find better care. Lyndhurst, NJ: Vitals; 2016. URL: http://www.vitals.com [accessed 2016-12-12] [WebCite Cache ID 6mh9Nl4je]
24. UCompareHealthCare: search and compare for better care. Marlborough, MA: UCompare Holdings, LLC; 2016. URL: http://www.ucomparehealthcare.com [accessed 2016-12-11] [WebCite Cache ID 6mh9ci3Ki]
25. RateMDs: doctors you can trust. San Jose, CA: RateMDs Inc; 2016. URL: https://www.ratemds.com/co/colorado-springs/ [accessed 2016-12-12] [WebCite Cache ID 6mh9geE5j]
26. Rao M, Clarke A, Sanderson C, Hammersley R. Patients' own assessments of quality of primary care compared with objective records based measures of technical quality of care: cross sectional study. BMJ 2006 Jul 1;333(7557):19 [FREE Full text] [doi: 10.1136/bmj.38874.499167.7C] [Medline: 16793783]
27. Chang JT, Hays RD, Shekelle PG, MacLean CH, Solomon DH, Reuben DB, et al. Patients' global ratings of their health care are not associated with the technical quality of their care. Ann Intern Med 2006 May 2;144(9):665-672. [Medline: 16670136]
28. Brown DL, Clarke S, Oakley J. Cardiac surgeon report cards, referral for cardiac surgery, and the ethical responsibilities of cardiologists. J Am Coll Cardiol 2012 Jun 19;59(25):2378-2382. [doi: 10.1016/j.jacc.2011.11.072] [Medline: 22698492]
29. Johnson C. Survey finds physicians very wary of doctor ratings. Physician Exec 2013;39(1):6-8, 10, 12. [Medline: 23437751]
30. Bleustein C, Rothschild DB, Valen A, Valatis E, Schweitzer L, Jones R. Wait times, patient satisfaction scores, and the perception of care. Am J Manag Care 2014 May;20(5):393-400. [Medline: 25181568]
31. Uhas AA, Camacho FT, Feldman SR, Balkrishnan R. The relationship between physician friendliness and caring, and patient satisfaction: findings from an internet-based survey. Patient 2008 Apr 1;1(2):91-96. [Medline: 22272805]

Abbreviations

CABG: coronary artery bypass graft

Edited by G Eysenbach; submitted 04.09.16; peer-reviewed by D Hanauer, S Atkinson; comments to author 29.09.16; revised version received 04.10.16; accepted 24.10.16; published 13.12.16.

Please cite as:
Okike K, Peter-Bibb TK, Xie KC, Okike ON
Association Between Physician Online Rating and Quality of Care
J Med Internet Res 2016;18(12):e324
URL: http://www.jmir.org/2016/12/e324/
doi:10.2196/jmir.6612
PMID:27965191

Viewpoint

IDEAS (Integrate, Design, Assess, and Share): A Framework and Toolkit of Strategies for the Development of More Effective Digital Interventions to Change Health Behavior

Sarah Ann Mummah[1,2], MPhil; Thomas N Robinson[1,3], MPH, MD; Abby C King[1,4], PhD; Christopher D Gardner[1*], PhD; Stephen Sutton[2*], PhD

[1]Stanford Prevention Research Center, Department of Medicine, Stanford University School of Medicine, Stanford, CA, United States
[2]Behavioural Science Group, Institute of Public Health, University of Cambridge, Cambridge, United Kingdom
[3]Stanford Solutions Science Lab, Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, United States
[4]Division of Epidemiology, Department of Health Research & Policy, Stanford, CA, United States
[*]these authors contributed equally

Corresponding Author:
Sarah Ann Mummah, MPhil
Stanford Prevention Research Center
Department of Medicine
Stanford University School of Medicine
1265 Welch Road
Stanford, CA, 94305-5411
United States
Phone: 1 650 723 7822
Fax: 1 650 725 6247
Email: sm885@cam.ac.uk

Abstract

Developing effective digital interventions to change health behavior has been a challenging goal for academics and industry players alike. Guiding intervention design using the best combination of approaches available is necessary if effective technologies are to be developed. Behavioral theory, design thinking, user-centered design, rigorous evaluation, and dissemination each have widely acknowledged merits in their application to digital health interventions. This paper introduces IDEAS, a step-by-step process for integrating these approaches to guide the development and evaluation of more effective digital interventions. IDEAS is comprised of 10 phases (empathize, specify, ground, ideate, prototype, gather, build, pilot, evaluate, and share), grouped into 4 overarching stages: Integrate, Design, Assess, and Share (IDEAS). Each of these phases is described and a summary of theory-based behavioral strategies that may inform intervention design is provided. The IDEAS framework strives to provide sufficient detail without being overly prescriptive so that it may be useful and readily applied by both investigators and industry partners in the development of their own mHealth, eHealth, and other digital health behavior change interventions.

(J Med Internet Res 2016;18(12):e317) doi:10.2196/jmir.5927

KEYWORDS

health behavior; design thinking; user-centered design; behavioral theory; behavior change techniques; digital interventions; mobile phones; digital health; telemedicine; diet; exercise; weight loss; smoking cessation; medication adherence; sleep; obesity

Introduction

Digital technology has rapidly and dramatically shifted how humans interact with the world and presents an unprecedented opportunity to develop, test, and widely disseminate effective health behavior change interventions. The prospect of modifying lifestyle behaviors, such as diet, physical activity, smoking, and sleep, to improve health outcomes has increasingly driven efforts in both academia and industry. However, achieving meaningful and sustained improvements in health behaviors has eluded researchers and industry players alike, and numerous challenges remain. Within industry, most digital health interventions are yet to incorporate theory-based strategies known to drive changes in health behaviors or undergo systematic testing to demonstrate their effectiveness [1-3]. Moreover, interventions are often plagued by rapidly declining retention rates [4], with a quarter of downloaded health apps used only once and three-quarters discontinued after the tenth use [5]. Within academia, interventions are more often grounded in behavioral theory and tested for their efficacy [6]. However, they face similar challenges around declining retention rates, and the rapid pace of new technology development makes it increasingly difficult for researchers to develop, pilot-test, and evaluate interventions before such technologies become outdated or obsolete [7,8]. In addition, researcher-driven technologies often do not benefit from numerous rapid cycles of fine-tuning based on user feedback nor do they usually become widely disseminated among the broader public [9].

In light of these challenges, investigators have called for digital health interventions to be (1) grounded in behavioral theory [1-3], (2) grounded in an in-depth qualitative understanding of the target population [10], (3) rapidly and iteratively developed with multiple stages of user feedback [11,12], (4) subject to rigorous evaluation [8,13], and (5) widely disseminated [14]. However, published frameworks to guide the development of such technologies are disparate and no single approach integrates these elements. In addition, specific recommendations for integrating principles from behavioral theory are lacking in available frameworks. These gaps limit our ability to advance health behavior change research and practice. As our knowledge evolves in this young field, investigators have been called on to publish the methods they use to develop interventions to further advance the field [15]. This paper responds to that call and aims to build on prior models by combining behavioral theory, design thinking, and intervention evaluation and dissemination through a systematic framework to guide the development of more effective digital health interventions. It introduces Integrate, Design, Assess, and Share (IDEAS), an integrated 10-phase process.

Essential Components of a Framework

Behavioral Theory

To maximize the potential efficacy of interventions to change health behavior, those who design interventions should have an understanding of theory or the hypothesized mechanisms underlying human behavior and behavior change [16]. Behavioral theory is widely acknowledged to be critical to the development of interventions to change health behavior [1-3,17-19] and increasing evidence suggests that health interventions grounded in theory are more effective than those without such theoretical foundations [19]. Among the most frequently used theories in health behavior research are social cognitive theory, the health belief model, and the transtheoretical model [20,21]. One approach to intervention development is to use one or more theories to identify the key constructs to be targeted in the intervention. For example, the health belief model suggests that a behavior change intervention should target perceived susceptibility, severity, benefits, and barriers. An alternative approach is to select behavior change techniques (eg, from those listed in the taxonomy of behavior change techniques [BCTs] [22]) and use these to construct the intervention. In principle, the two approaches could be used in combination, selecting techniques that are believed based on theory and/or evidence to be likely to produce change in the targeted behaviors.

Design Thinking

A recent consensus statement on the prevention of noncommunicable diseases emphasized the importance of using human-centered design, or *design thinking*, to develop effective and innovative interventions [23]. First coined by David Kelley, a founder of the design firm IDEO and the Stanford University Hasso Plattner Institute of Design (d.school), design thinking is intended to guide the development of more creative and innovative solutions [24]. The process has been summarized by the Stanford d.school as encompassing 5 phases: "empathize" (understand target population), "define" (identify goals and scope), "ideate" (brainstorm potential solutions), "prototype" (mock up primitive potential solutions), and "test" (gather feedback from target users) [25]. Design thinking starts by reframing the context for behavioral change around "what matters most" to a target group rather than "what's the matter" with them [23]. This approach enables the design of more empathetic solutions that are more desirable to target populations. The importance of grounding behavioral health interventions in a deep understanding of a target population has been emphasized by experts in public health and health psychology [10,23]. Design thinking next involves rapidly and iteratively brainstorming, prototyping, and gathering user feedback on potential solutions. This process is similar to *user-centered design*, first presented by Norman and Draper [26], and involves gathering user feedback throughout intervention development, refining designs through prototyping and iteration, and including multidisciplinary skills and perspectives [27]. Core to design thinking is the notion that everyone has the potential to be highly creative and can learn to apply design-based approaches such as need-finding, brainstorming, prototyping, and iteration to unlock their creative potential [28]. It has been suggested that design thinking may increase the self-efficacy of those who practice the process [29] and frees even the most novice designers to generate more creative ideas and solutions. As a result, the process has been increasingly adopted among a diverse range of academics and industry players [24] and has been adapted to better serve the varied contexts in which it is used, particularly through discipline-specific frameworks and processes [30,31].

Evaluation and Dissemination

There have long been calls for the evaluation and dissemination of digital health interventions [8]. Rigorous evaluation is essential to judge whether or not an intervention achieves its desired effect. Published evaluations also contribute to the literature and evidence base, helping uncover the potential and challenges of using digital interventions to improve health outcomes [8,13]. Rigorous evaluation methods also enable more reliable conclusions. It has been recommended that the purpose of evaluation, the balance of potential benefits and risks, and the level of resource investment required by providers and users can help drive the necessary level of evidence needed [8]. As the UK Medical Research Council emphasizes, evaluation efforts should employ randomization whenever possible because it is considered the most robust approach to preventing several forms of bias. Without randomization to the intervention or a control or comparison condition, it is not possible to determine whether the intervention itself was responsible for the observed effects or whether a selected group of participants, for instance, happened to be highly motivated and might have improved without the intervention [19].

Moreover, dissemination is crucial if digital health interventions are to fulfill their potential. One of the great promises of digital interventions is their ability to reach broad segments of the population with minimal cost [14]. However, most publicly available apps to promote dietary behavior change have not been assessed in rigorous randomized controlled efficacy trials [32,33]. In turn, most digital health interventions that have demonstrated efficacy in peer-reviewed trials are not available to the public; instead, they have been created ad hoc for research purposes [32]. Thus, interventions that are publicly available have not been evaluated and those that have been evaluated are not publicly available [32]. As a result, there is a need for effective interventions to be more widely disseminated to populations that may benefit.

Existing Frameworks and Limitations

Using the best combination of recommended approaches to guide intervention design is important if effective technologies to change health behavior are to be developed. Researchers often rely on published frameworks to guide them through the process of designing digital interventions [34]. However, currently available frameworks are numerous, disparate, and do not fully integrate behavioral theory, design thinking, and evaluation and dissemination. Although many electronic health (eHealth) frameworks exist, most envision their objective as guiding the development of technologies to facilitate medical or patient care (eg, patient-physician communication, access to medical records) rather than to modify health behavior [35,36]. For instance, Van Velsen et al [36] propose a “requirements development” approach in which stakeholder interviews are meant to lead directly to a list of technical specifications to be developed (eg, one-stop portal for patient information). This approach may be appropriate for building products to facilitate logistics or care provision in medical settings, but designing for behavior change is a different type of endeavor that requires thoughtful integration of behavioral theory and evidence.

Relatively few frameworks focus on guiding the development of digital interventions for the express purpose of changing health behavior. Among the frameworks that do are Yardley et al’s [10] person-based approach, Ludden et al’s [37] design research perspective, and Brown et al’s [38] health information technology usability evaluation model. Although each of these approaches provides valuable guidance for investigators, each focuses on a particular aspect of intervention development and none provides guidance on behavioral strategies that may be used in intervention design. Hekler et al’s [12] process, referred to as behavioral science-informed user experience design, combines a user-centered design approach with the use of behavioral theory-driven strategies. This process notably suggests the integration of user-centered and theory-based approaches, but it does not provide step-by-step guidance on how others may replicate the approach. Whittaker and colleagues [15] have proposed perhaps the most comprehensive step-by-step framework to date which involves 5 phases (focus groups, pretesting, pilot, randomized controlled trial [RCT], interviews) to guide overall mobile health intervention development and evaluation. Whittaker’s framework has been applied to the iterative development of numerous mobile health interventions [39-42], includes stages for user feedback and evaluation, and states the importance of using behavioral theory. However, it does not make use of design thinking approaches such as ideation, brainstorming, or rapid prototyping, nor does it include specific guidance on behavioral strategies that may inform intervention design. Although it has been suggested that design thinking and behavioral science can together inform the development of more effective digital health interventions [12], no published frameworks appear to combine behavioral theory, design thinking, and evaluation and dissemination into a comprehensive step-by-step process for guiding digital interventions to change health behavior. This gap limits our ability to advance health behavior change research and practice.

IDEAS Framework

Overview

To address the need for a framework that more fully integrates strengths from behavioral theory, design thinking, and evaluation and dissemination, we introduce IDEAS, a framework to better guide the development of digital health interventions to change behavior. IDEAS was informed by a multisector team of researchers, designers, and engineers, and was then applied to and refined in the iterative development of Vegethon, a mobile health (mHealth) intervention that demonstrated user acceptability and initial efficacy [43]. IDEAS consists of 10 phases: (1) empathize with target users, (2) specify target behavior, (3) ground in behavioral theory, (4) ideate implementation strategies, (5) prototype potential products, (6) gather user feedback, (7) build a minimum viable product, (8) pilot test to assess potential efficacy and usability, (9) evaluate efficacy in an RCT, and (10) share intervention and findings. These phases are grouped into 4 overarching categories: Integrate, Design, Assess, and Share (Figure 1).

Figure 1. IDEAS (Integrate, Design, Assess, and Share) framework for developing digital health behavior change interventions.

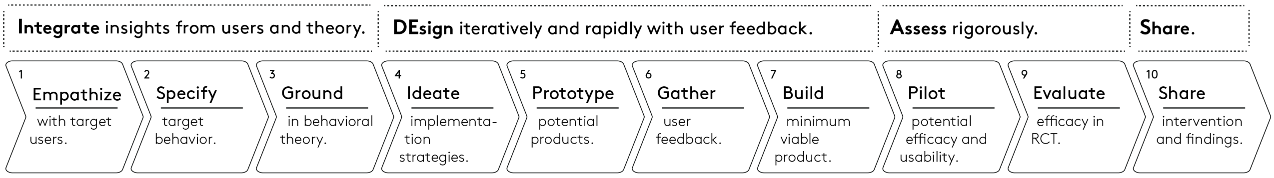
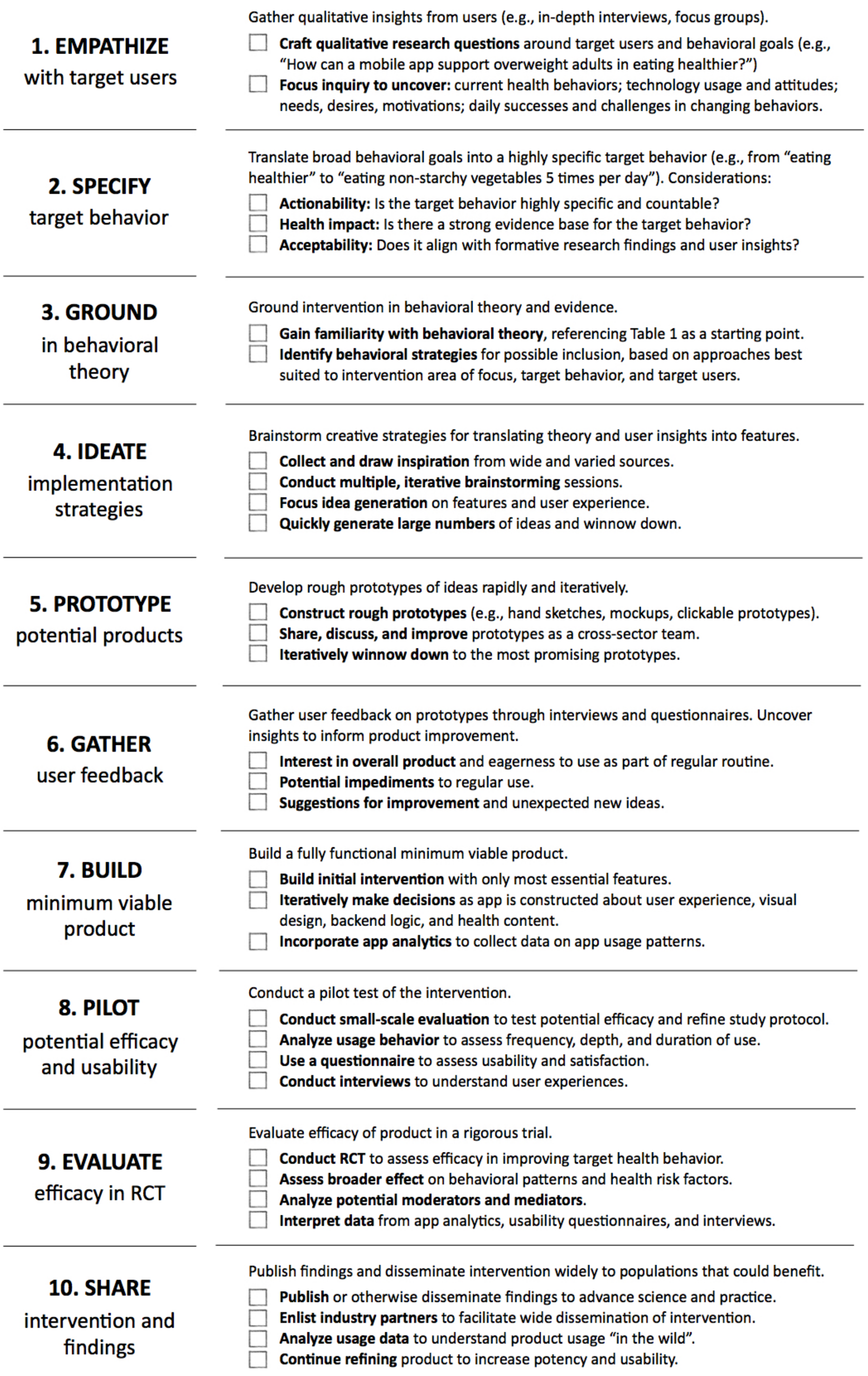


Figure 2. IDEAS (Integrate, Design, Assess, and Share) phases for developing digital health behavior change interventions.



1. EMPATHIZE with target users

Gather qualitative insights from users (e.g., in-depth interviews, focus groups).

☐ **Craft qualitative research questions** around target users and behavioral goals (e.g., “How can a mobile app support overweight adults in eating healthier?”)

☐ **Focus inquiry to uncover:** current health behaviors; technology usage and attitudes; needs, desires, motivations; daily successes and challenges in changing behaviors.

2. SPECIFY target behavior

Translate broad behavioral goals into a highly specific target behavior (e.g., from “eating healthier” to “eating non-starchy vegetables 5 times per day”). Considerations:

☐ **Actionability:** Is the target behavior highly specific and countable?

☐ **Health impact:** Is there a strong evidence base for the target behavior?

☐ **Acceptability:** Does it align with formative research findings and user insights?

3. GROUND in behavioral theory

Ground intervention in behavioral theory and evidence.

☐ **Gain familiarity with behavioral theory**, referencing Table 1 as a starting point.

☐ **Identify behavioral strategies** for possible inclusion, based on approaches best suited to intervention area of focus, target behavior, and target users.

4. IDEATE implementation strategies

Brainstorm creative strategies for translating theory and user insights into features.

☐ **Collect and draw inspiration** from wide and varied sources.

☐ **Conduct multiple, iterative brainstorming** sessions.

☐ **Focus idea generation** on features and user experience.

☐ **Quickly generate large numbers** of ideas and winnow down.

5. PROTOTYPE potential products

Develop rough prototypes of ideas rapidly and iteratively.

☐ **Construct rough prototypes** (e.g., hand sketches, mockups, clickable prototypes).

☐ **Share, discuss, and improve** prototypes as a cross-sector team.

☐ **Iteratively winnow down** to the most promising prototypes.

6. GATHER user feedback

Gather user feedback on prototypes through interviews and questionnaires. Uncover insights to inform product improvement.

☐ **Interest in overall product** and eagerness to use as part of regular routine.

☐ **Potential impediments** to regular use.

☐ **Suggestions for improvement** and unexpected new ideas.

7. BUILD minimum viable product

Build a fully functional minimum viable product.

☐ **Build initial intervention** with only most essential features.

☐ **Iteratively make decisions** as app is constructed about user experience, visual design, backend logic, and health content.

☐ **Incorporate app analytics** to collect data on app usage patterns.

8. PILOT potential efficacy and usability

Conduct a pilot test of the intervention.

☐ **Conduct small-scale evaluation** to test potential efficacy and refine study protocol.

☐ **Analyze usage behavior** to assess frequency, depth, and duration of use.

☐ **Use a questionnaire** to assess usability and satisfaction.

☐ **Conduct interviews** to understand user experiences.

9. EVALUATE efficacy in RCT

Evaluate efficacy of product in a rigorous trial.

☐ **Conduct RCT** to assess efficacy in improving target health behavior.

☐ **Assess broader effect** on behavioral patterns and health risk factors.

☐ **Analyze potential moderators and mediators.**

☐ **Interpret data** from app analytics, usability questionnaires, and interviews.

10. SHARE intervention and findings

Publish findings and disseminate intervention widely to populations that could benefit.

☐ **Publish** or otherwise disseminate findings to advance science and practice.

☐ **Enlist industry partners** to facilitate wide dissemination of intervention.

☐ **Analyze usage data** to understand product usage “in the wild”.

☐ **Continue refining** product to increase potency and usability.

IDEAS aims to facilitate the creation of more effective interventions by using strengths from a combination of disciplines at the intersection of digital health. The framework builds on design thinking and user-centered design in that it is iterative, engages multidisciplinary perspectives through a cross-sector team, includes stages for ideation and prototyping (phases 4, 5, 7), and integrates user insights throughout (phases 1 and 6). It focuses the design process around defining behavioral goals grounded in evidence (phase 2) and encourages the inclusion of theory-driven behavioral strategies both initially (phase 3) and throughout the design process. IDEAS integrates evaluation methods captured by the UK Medical Research Council framework [19] by emphasizing the importance of rigorously evaluating behavioral outcomes through both pilot and more substantially powered RCTs (phases 8 and 9). Finally, given the abundance of low-quality interventions currently available to the public, IDEAS concludes with a stage for dissemination (phase 10), which capitalizes on the readily scalable nature of digital interventions to provide access to target populations that may benefit. Dissemination also includes sharing findings with other audiences that can use the findings to advance the field and science of behavior change.

The 10 IDEAS phases are summarized in a step-by-step fashion in Figure 2. As with design thinking, although the IDEAS phases are presented sequentially, they are not necessarily intended to be carried out linearly, and projects are encouraged to loop back through the phases in an iterative fashion as different ideas and directions are explored and refined [44].

Phase 1: Empathize With Target Users

The first 3 IDEAS phases (empathize, specify, and ground) aim to integrate insights from users and behavioral theory and orient the intervention development process around a measurable target behavior. These phases are primarily about information gathering to help inform the development in later stages of a more well-accepted, theory-driven intervention.

In phase 1, qualitative research is undertaken to gain a deeper understanding of the selected target population and their needs [10,45]. This need-finding stage can include observations, interviews, focus groups, and questionnaires [46]. Designers gain insight into users' unmet or latent needs, which they may not necessarily be aware of or be able to articulate [47]. The goal of this stage is to move beyond identifying incremental improvements that users might be able to articulate (eg, refinements to a health app they already enjoy using) and to instead uncover deeper needs, values, and motivations that may help inspire more innovative and creative solutions [44]. For instance, nuanced insights, such as not wanting to feel like a failure or lacking family support for healthier cooking, while not directly suggestive of potential intervention solutions, may equip team members with the background and context necessary to develop an intervention that helps users feel supported with a virtual social support group and positive language. By engaging all members of the intervention development team in this qualitative research stage, it may be possible to guide the development of solutions that are more relevant and acceptable to the target population [48].

Phase 2: Specify Target Behavior

Insights gathered from users help inform the next phase, in which a highly specific target behavior is selected. This target behavior defines both the purpose of the intervention as well as the outcomes by which the intervention will be judged. For instance, an initially broad intervention goal of "increase physical activity" may be refined to "take 10,000 steps each day." The narrowing of the intervention goal in this way helps to focus the scope of idea generation and has been associated with success in the context of highly innovative concepts [49].

Multiple, and sometimes competing, factors may inform the selection of the target behavior. Insights from users in the previous phase may help investigators identify appropriate potential target behaviors that would be well accepted by individuals. A review of the literature may be conducted to identify the health benefits of possible target behaviors under consideration. For example, before deciding whether to target greater vegetable consumption or greater breakfast consumption, it would be advantageous to understand whether one behavior is more likely to have a greater health impact than the other. Some behaviors may be more susceptible to change than others, particularly if they are readily countable and therefore easier for users to self-monitor and modify. Some behaviors may also have the potential to produce a beneficial cascade effect, improving other health behaviors or aspects of the same health behavior [50]. Qualitative investigations have also shown that users may be more interested in technologies that frame behaviors as actions to increase as opposed to decrease (eg, increasing the number of days per week free of sugar-sweetened beverages vs cutting back on sugar-sweetened beverages) [51]. Thus, principles that may guide the selection of the intervention target behavior include (1) a behavior that is well accepted by the target population, (2) an evidence base demonstrating a significant health benefit to changing the target behavior, (3) a behavior that is highly specific and countable and therefore more actionable, (4) a behavior that is framed as something to increase versus decrease, and (5) optimally a behavior that if changed has the potential for producing a beneficial cascade effect on related and reinforcing behaviors.

Phase 3: Ground in Behavioral Theory

Once a target behavior is identified, strategies are explored to ground intervention development in behavioral theory. The intervention design team may seek to gain familiarity with a range of theory-driven behavioral strategies available for inclusion. Strategies best suited to the target behavior and intervention delivery medium are identified for possible inclusion. Researchers may draw from behavioral theories, such as social cognitive theory [52], which are frequently used in academia to guide mobile health interventions [2]. Researchers may also draw from collections of theory-driven behavioral strategies, which have increasingly been introduced by investigators. For example, Michie et al's [22] taxonomy summarizes 93 behavior change techniques such as goal setting, self-monitoring, feedback, prompts/cues, and action planning.

In our cumulative experience, we have also found it useful to apply a theory-driven process motivation lens during intervention development [53]. Process motivators, first introduced by Robinson [53], may be used to make the process itself of behavior change more engaging and intrinsically rewarding. Process motivators stand in contrast to *outcome motivators*, which focus on the eventual outcomes of behavior change such as weight loss, physical appearance, becoming more fit, and reducing risks of future chronic diseases [53]. Although outcome motivators have historically dominated medical and public health interventions (eg, smoking cessation interventions emphasizing reduced risk of lung cancer), such rational appeals are limited in their motivational power to initiate and sustain behavior changes [54,55]. Outcome motivators rely on results that are often delayed and difficult to achieve and maintain, reducing their motivational impact and an individual's self-efficacy for behavior change [52,53,56]. By contrast, interventions that emphasize motivation for participating in the intervention itself, or the process of behavior change, may be more effective in initiating and sustaining behavior changes [53]. It has been suggested that process motivators, such as fun, taste, pride, choice, and challenge, can make the process of eating healthier or of engaging in physical activity more inherently enjoyable and desirable [53]. Table 1 presents examples of behavioral strategies using process motivation, adapted and extended from Robinson [53], and based on research on intrinsic motivation and interventions to change behavior [53,57-62], which may be useful to investigators in the development of their own interventions.

Table 1. Toolkit of behavioral strategies using process motivation to guide intervention design.

Behavioral strategy	Description
Challenge	Maintain optimal levels of moderate challenge (ie, not too hard, not too easy)
Choice/control	Provide objective and perceived choice and control over one’s environment and actions
Community	Provide social meaning (public recognition, identification with desirable group) for accomplishments
Competence	Provide immediate, frequent, clear, constructive, encouraging positive feedback following success
Competition	Facilitate social comparison and competition among individuals, groups, or teams
Context	Embed intervention into real/imaginary contexts with stories/characters
Curiosity	Provide sensory (color, taste, sound) and cognitive (mystery) curiosity and surprise
Growth mindset	Cultivate belief that behaviors/preferences (eg, for foods, activity levels) are malleable with effort
Identity	Facilitate an identity shift related to the behavior change (eg, someone who is now a runner)
Personalization	Personalize intervention using an individual’s name and personally relevant content
Pride	Cultivate pride and a sense of accomplishment
Piggybacking	Engage individuals in social movements (eg, animal rights) to harness deeper values
Reframing	Cast the purpose of a behavior in a more positive light to improve thoughts or feelings about it
Taste	Emphasize the taste and texture of healthier foods
Teamwork	Facilitate cooperation and teamwork among individuals, groups, or teams

Regardless of which behavioral theories or strategies are used or preferred by investigators, the goal of this phase is to consider ways in which behavioral theory can be incorporated into intervention development. Equipped with behavioral theory and strategies, subsequent stages enable the creative translation of such strategies into tangible and specific intervention components. Theory-driven insights can also be instrumental in helping guide the interpretation of user interview findings and addressing gaps in user understanding and awareness. At this stage, the team may wish to explore, discuss, compare, and reconcile insights from users and behavioral strategies in order to prepare for the brainstorming phase.

Phase 4: Ideate Creative Implementation Strategies

The next 4 phases (ideate, prototype, gather, and build) involve a highly iterative design process that focuses idea generation around the target behavior and takes insights from users and behavioral theory into account throughout.

To begin, a series of group brainstorming sessions is held to ideate potential intervention solutions, components, and features. The agreed-upon target behavior focuses the scope of idea generation [25] and the aim is to generate a large number of diverse ideas that could influence the target behavior. Previously gathered user insights and behavioral theories help inspire the range of ideas. For instance, a “self-monitoring” behavioral strategy in the context of a target behavior to “increase daily steps” might inspire an idea to engage users in taking a photograph of the outdoor scenery each time they go for a run. User insights suggesting that individuals prefer outdoor running to other types of activities might lend further credibility to such an idea. Hundreds of ideas may be generated “ranging from the absurd to the obvious” (p 31 [44]). Adhesive notes can be used to capture different ideas (one idea per note), which can be displayed on a wall or board, and visual representations of ideas are encouraged to facilitate communication of ideas to other team members [44]. Divergent thinking is achieved through interdisciplinary teams engaging in structured brainstorming sessions. Team members are encouraged to generate ideas that build on prior ideas and/or are divergent from those already suggested. Deferring judgment is a central rule during this phase to encourage rather than discourage idea generation [63]. Throughout the process, weaker ideas drop off early on, whereas stronger ideas “naturally rise to the top” [44]. This phase is ideally undertaken with a cross-sector team to enable the generation of greater numbers of more varied and creative ideas. As brainstorming progresses, discussing and debating ideas during this highly creative stage is recommended because it has been associated with more novel innovations and produces a more thorough exploration of possible solutions [64].

Phase 5: Prototype Potential Products

Prototypes are then created to share and discuss ideas with team members and to facilitate both further ideation (phase 4) and the rapid gathering of user feedback (phase 6). The prototyping phase is exemplified by the adage “enlightened trial and error outperforms the planning of flawless intellect” (p 1 [65]). Prototypes may be sketched, hand crafted using primitive materials, and/or developed into clickable mockups of digital interfaces. In the earliest stages, low-cost, low-fidelity (ie, “quick and dirty”) prototypes are developed rapidly to gather feedback from users early and often. This approach allows intervention designers to quickly and cheaply gather feedback on many different possible intervention approaches before investing significant resources in any one particular approach or suboptimal solution [25,66]. Experimental studies have demonstrated that the act of developing multiple prototypes in parallel (vs sequentially) leads to objectively stronger results [65]. In one experiment, when novice designers were instructed to develop multiple prototypes (vs a single prototype) before obtaining user feedback, designs overall tended to be more divergent and final prototypes were superior as measured by click-through data and blinded expert ratings. These findings were explained by qualitative data suggesting that parallel prototyping reduced fixation on a particular idea and encouraged designers to instead explore multiple directions before optimizing in any one direction. By contrast, sequential prototyping implicitly encouraged the refinement of the initial prototype at the expense of exploration of more divergent alternatives [29]. Although time constraints often lead teams to focus on the realization of a single idea rather than the iteration of multiple ideas [67], prototype iteration has been shown to lead to objectively stronger final products even when a team is under time constraints [65].

Phase 6: Gather User Feedback

User feedback is then gathered on prototypes. Methods may include informal conversations, usability tests, surveys, and in-depth qualitative interviews. For instance, during a usability test of an early prototype, a researcher may observe participants use the intervention and participants may be asked to think aloud by providing commentary during the process. These methods of inquiry seek to uncover users' interests in the overall product and eagerness to use it as part of their typical routine, potential impediments to usage on a regular basis, suggestions for improvement, and unexpected new ideas or opportunities. When this phase is initially employed in the design process, the focus is on quickly gathering initial user impressions and inspiring further divergent ideation. As this phase is repeated in quick succession with the ideate and prototype phases such that the phases inform one another [68], the goal is increasingly to inform concept refinement and focus. Thus, the ideate, prototype, and gather phases are carried out in a fluid and iterative fashion until a more refined design solution is reached [25].

Phase 7: Build a Minimum Viable Product

Next, a fully functional minimum viable product (MVP) is developed to facilitate early learning from users. This stage focuses on a level of detail beyond the prototype stage with decisions made regarding user experience, visual design, content, and logic (eg, to calculate graphs, to trigger push notifications). Analytics are built into the product to enable the collection of a rich dataset capturing patterns of usage. The use of analytics such as metadata may be applied in the development of just-in-time, adaptive interventions (JITAIs), in which real-time data are used to tailor an intervention to the dynamically changing needs of a user [69,70]. The practical challenges and realities of building a digital intervention surface numerous decisions that need to be made with the technical developers. A multisector team can include perspectives to help ensure that the influence of behavioral theory, user needs and desires, and technical, financial, and practical feasibility are properly balanced and reconciled throughout the decision-making process. A core tenet of the MVP concept is to avoid wasting precious time and resources on perfecting a product that may be substantially altered in subsequent phases based on user feedback [71]. The goal, therefore, is to quickly develop only a minimum version of the product necessary to facilitate pilot testing among users, without incorporating additional unnecessary features. However, this industry-driven tenet must be balanced with the common assumption among behavioral scientists that many behavioral strategies are most effective when used in concert, rather than in isolation [39], which may encourage the inclusion of more behavioral strategies. Weighing these opposing priorities (ie, fewer features vs more theory-driven strategies) will determine how theory- and feature-rich the minimum viable intervention becomes at this stage. The degree to which an intervention is theory- and feature-rich will vary in each case and depend on the unique resource (eg, financial, human, time) constraints as well as the behavioral needs of the intervention. As with all phases in IDEAS, this phase can be conducted iteratively with other phases, particularly phase 6, to gather user feedback to help inform product decisions before pilot testing.

Phase 8: Pilot Test

The final 3 phases (pilot, evaluate, and share) aim to assess and disseminate the final intervention and evaluation results. This process begins by subjecting the MVP to pilot testing to gather information on the usability, feasibility, and potential efficacy of the first viable version of the intervention [8,15,39]. For example, usability and satisfaction questionnaires may ask individuals on a 5-point scale how strongly they agree with statements such as "I have found this app easy to use," "this product has motivated me to be more physically active," or "I have had a hard time remembering to use this product" [72]. Analysis of product usage data helps identify usage patterns and peak usage times, which may inform notification timing. Qualitative and usability interviews lend insight into preferred features, problematic or confusing user flows, and opportunities for intervention modifications or improvements. Study design and implementation procedures, such as recruitment, enrollment, retention, and data collection, are tested and refined to avoid problems in subsequent larger-scale evaluation studies [73]. Randomization, which may [74] or may not [15] be used at this pilot stage, enables the potential efficacy of the intervention to be assessed [75]. Based on the wealth of information gathered during this stage, the intervention is further refined until it is ready for a larger-scale evaluation.

Phase 9: Evaluate Efficacy

A sufficiently powered RCT is then conducted to test intervention efficacy. RCTs are considered the most robust study design for evaluating interventions due to their ability to minimize bias [8,19,76]. Outcomes may include measures of the original target health behavior; mediators, proximal behaviors, or conditions that may lie on the causal pathway to behavior change; and broader effects on related health behaviors and risk factors. Assessing possible moderators, or preexisting factors that may help define characteristics of users that respond more or less to the intervention, has also been recommended to better define the appropriate target audience for ultimate dissemination [77]. Given the rapidly evolving nature of mobile technology and the time involved in carrying out a rigorous evaluation study, the digital intervention may need to be continuously refined during the trial to ensure that it does not become obsolete by the time the trial is completed [13]. Researchers need to weigh the advantages of such an approach with the potential limitations to internal validity that are introduced if the intervention is modified too dramatically. To yield reliable and valid results, several best practices for conducting an RCT are followed. Randomization is used to eliminate conscious and unconscious selection bias in the assignment of participants to the intervention versus a control condition [78]. Proper randomization is achieved when the allocation sequence is unpredictable and concealed, such that the researcher enrolling participants does not know in advance which treatment the next participant will receive [79,80]. The widely recommended "intention-to-treat" analysis approach is used to preserve the benefits of randomization, in which all randomized participants are included in the analysis and retained in the original groups to which they were allocated, regardless of adherence to the digital intervention, missing data, or dropouts [81,82]. Data collectors and researchers who perform statistical analyses are blinded to intervention assignment to prevent the introduction of biases in the interpretation or analysis of data. For further guidance on RCTs, researchers may consult the evidence-based Consolidated Standards of Reporting Trials (CONSORT) 2010 guidelines [83].

Phase 10: Share Widely

There are two types of sharing implied in the IDEAS framework. First, learnings and evaluation results should be disseminated to both researchers and digital health intervention developers to help advance the field and improve the effectiveness of future interventions. For example, the evaluation trial's findings should be published, including a description of the final intervention and its theoretical basis and/or use of behavior change techniques, such as those using standardized language and descriptors from a taxonomy of behavior change techniques [22]. The second type of sharing involves disseminating efficacious interventions more broadly with appropriate target audiences. One of the greatest promises of digital health interventions is their potential to be widely scaled to users at a relatively low cost [84]. Thus, once intervention efficacy has been demonstrated, the intervention is disseminated to the broader population that may benefit the most, as demonstrated in the evaluation. Scaling an intervention for wide dissemination is not a trivial undertaking. Health-related intervention goals may be less salient in certain dissemination settings where, for instance, health care providers, health insurers, and/or employers may be more motivated by cost savings. The systems architecture (eg, servers, databases, storage systems) of an intervention may need to be adjusted to account for greater traffic at scale. Further intervention refinements may be warranted based on dissemination goals [10]. Cost considerations, including marketing, maintenance, and ongoing implementation, may significantly influence adoption [14]. Academic researchers may wish to form strategic partnerships with relevant public and private sector organizations to support the more effective and sustainable dissemination of the final intervention [84].

How to Apply IDEAS

Engage a Multidisciplinary Team

Digital health is increasingly acknowledged to be an inherently transdisciplinary endeavor in which involving user-centered designers, programmers, behavioral researchers (particularly those working in the area of behavioral intervention development), and industry partners is crucial [9,11]. The effectiveness of digital health interventions rests on multiple diverse factors including esthetic design, behavioral theory, evidence grounding, user centeredness, technical capacity, and demonstrated efficacy. Only a true multidisciplinary team will have the knowledge and expertise to address these complex factors in tandem [27]. Engaging all team members throughout the full intervention development process may help the team proceed through each of the phases iteratively, nimbly, collaboratively, and with greater buy-in at all stages. This practice can also help ensure that all team members understand users' needs and the theoretical grounding underpinning the intervention. Finally, enlisting the full multidisciplinary team may facilitate the generation of more highly divergent ideas and prototypes, which has been shown to lead to more creative and efficacious final intervention designs [64].

Use a Flexible, Iterative Approach

As with design thinking, the teams that stand to benefit the most from applying the IDEAS framework are those who will apply the stages flexibly, pursuing multiple stages in parallel, in combination, more than once, and/or iteratively, as new directions are explored and refined [68]. For example, some high-performing teams have been shown to combine, in particular, the ideate, prototype, and gather phases, rapidly shifting between them as needed. A single brainstorming session (ideate phase) may flow directly into the use of highly primitive prototypes (prototype phase) to communicate ideas to team members, generate new ideas, and/or gather user feedback (gather phase) to facilitate further brainstorming (ideate phase) [68]. In this way, the 10 IDEAS phases are more like "a system of overlapping spaces, rather than a sequence of orderly steps" as design thinking has been described [44].

Transition Between Divergence and Convergence

A key consideration in applying the framework involves balancing the opposing processes of divergent and convergent thinking. Idea divergence, or the generation of highly diverse and varied ideas, is critical during early stages of the intervention development process to fully consider the range of user experiences and possible solutions. However, idea convergence, or the narrowing of possible ideas, becomes equally important to identify a target behavior and refine potential solutions without being hampered by a continual revision of prior decisions which may impede progress. High-performing teams have been shown to shift their behavior nimbly and repeatedly as needed throughout the design process, exhibiting greater debate and divergence during "concept generation" stages and less debate and more focused attention during "concept selection" stages [68]. In a space where digital technology is young and the best solutions are likely yet to be uncovered, a highly divergent approach in brainstorming possible solutions may maximize the likelihood of identifying the most potent solutions. As the design process progresses, successful teams may find it advantageous to transition from divergence to convergence, particularly once a set of possible solutions have been identified for refinement [85].

Discussion

Overview

This paper introduces IDEAS, a framework to guide the design, development, and evaluation of digital interventions to change health behavior. It includes a summary of behavioral strategies that may be useful for intervention developers seeking to apply IDEAS in the development of their own interventions. IDEAS is among the few frameworks that aim to guide the development of digital interventions to change health behavior. It builds on Whittaker et al's [15] framework by incorporating phases for ideation and prototyping that may contribute to more creative interventions and by providing guidance on behavioral strategies that may inform more effective technologies. It draws on strengths from Yardley et al's [10] person-based approach through a focus on qualitative inquiry and Hekler et al's [12] behavioral science-informed user experience design model through a focus on behavioral strategies and user-centered design. Although the importance of combining behavioral theory and design thinking has been emphasized [12], IDEAS appears to be the first framework to provide step-by-step guidance on integrating these approaches.

Strengths

Among the strengths of IDEAS is its provision of a toolkit of behavioral strategies grounded in process motivation, which may aid and inspire intervention design. These strategies include theory-driven approaches not presently captured by the taxonomy of behavior change techniques [22] and may serve as a useful additional resource for intervention developers seeking to develop their own solutions. Moreover, theory by nature is abstract [21] and intervention developers may not know how to translate such insights into concrete intervention features. The IDEAS framework fills this gap by suggesting that ideating, prototyping, and gathering user feedback may be used iteratively and in quick succession to incrementally translate theories into highly relevant and practical intervention components. IDEAS also addresses the topic of creativity, which has been given considerably less attention in the mHealth and eHealth design literature [36]. Through structured brainstorming sessions and rapid prototyping with user feedback, intervention developers aiming to incorporate greater amounts of creativity into their designs are provided with practical guidance on how to do so. In addition, by integrating user insights throughout the intervention development process, IDEAS helps guard against an unexpected mismatch between intervention goals and user goals, which may lead to low user satisfaction and poor adherence to an intervention [37,86]. Finally, the promise of digital interventions lies in their potential for reaching broad segments of the population [14]; where other frameworks have neglected dissemination altogether, the inclusion of dissemination as a key phase may help more intervention developers take this stage into consideration to advance the field overall and produce greater population impacts on health.

Limitations

Several limitations to this framework exist. Despite research exploring the application of design-based approaches among novice designers [68], less-experienced users may find it difficult to effectively apply some of the suggested methods. There are practical challenges inherent to working with a multidisciplinary team and team members may disagree in a counterproductive manner or find it difficult to advance potential intervention solutions. The behavioral strategies presented are not exhaustive. The framework does not suggest the randomized evaluation of isolated intervention components [87]. However, it is commonly argued that behavioral strategies work best in concert with one another; thus, evaluating individual intervention components separately may not necessarily be advantageous in this context [39]. This framework is not the only approach to intervention development and many other useful approaches may complement or advance IDEAS. Future research is recommended to systematically evaluate the IDEAS framework.

Conclusions

The IDEAS framework is proposed to guide the design, development, and rigorous evaluation of more creative and effective digital health interventions. It integrates behavioral theory, design thinking and user-centered design, and evaluation and dissemination, and summarizes a list of theory-driven behavioral strategies that may be useful to intervention developers. IDEAS is intended to accelerate the translation of behavioral theory and evidence into industry practice where most digital health interventions are born. Other researchers who use IDEAS or alternative frameworks are encouraged to report on their processes and outcomes so that we and others may learn from their experiences and continue to improve the quality, efficacy, and effectiveness of our digital health interventions.

Acknowledgments

This research was supported by a Gates Cambridge Scholarship (SAM), Award Number R01DK0091831 from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) of the National Institutes of Health (NIH) (CDG), an unrestricted gift from the Nutrition Science Initiative (NuSI) (CDG), and a grant from the Stanford Child Health Research Institute (TNR). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDDK or NIH.

The authors would like to thank Katherine Dotter, Dustin Yoder, Jae Berman, Ryan Swanson, Maya Mathur, and Haley Schmoyer for helping inform the development of the framework.

Authors' Contributions

SAM developed the framework and wrote the manuscript. SS, TNR, CDG, and ACK contributed guidance and consultation throughout and provided feedback on the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
2. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
3. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]
4. Chen J, Cade JE, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. JMIR Mhealth Uhealth 2015;3(4):e104 [FREE Full text] [doi: 10.2196/mhealth.4334] [Medline: 26678569]
5. McLean V. Motivating patients to use smartphone health apps. PRWeb. 2011 Apr 25. URL: http://www.prweb.com/releases/2011/04/prweb5268884.htm [accessed 2016-06-26] [WebCite Cache ID 6iYtbn6pb]
6. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
7. Baker TB, Gustafson DH, Shah D. How can research keep up with eHealth? Ten strategies for increasing the timeliness and usefulness of eHealth research. J Med Internet Res 2014;16(2):e36 [FREE Full text] [doi: 10.2196/jmir.2925] [Medline: 24554442]
8. Robinson T, Patrick K, Eng T, Gustafson D. An evidence-based approach to interactive health communication: a challenge to medicine in the information age. Science Panel on Interactive Communication and Health. JAMA 1998;280(14):1264-1269.
9. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [FREE Full text] [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
10. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055]
11. Tate EB, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz MA, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. Transl Behav Med 2013 Dec;3(4):406-415 [FREE Full text] [doi: 10.1007/s13142-013-0222-3] [Medline: 24294329]
12. Hekler E, King A, Banerjee B, Robinson T, Alonso M, Cirimele J. A case study of BSUED: behavioral science-informed user experience design. In: Proceedings of the SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems. 2011 Presented at: CHI 2011; May 7–12, 2011; Vancouver, BC p. 7-12.
13. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
14. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
15. Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012 May;17 Suppl 1:11-21. [doi: 10.1080/10810730.2011.649103] [Medline: 22548594]
16. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol Rev 2015 Sep;9(3):323-344 [FREE Full text] [doi: 10.1080/17437199.2014.941722] [Medline: 25104107]
17. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696 [FREE Full text] [Medline: 10987780]

18. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. BMJ 2007 Mar 3;334(7591):455-459 [FREE Full text] [doi: 10.1136/bmj.39108.379965.BE] [Medline: 17332585]
19. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]
20. Painter JE, Borba CP, Hynes M, Mays D, Glanz K. The use of theory in health behavior research from 2000 to 2005: a systematic review. Ann Behav Med 2008 Jun;35(3):358-362. [doi: 10.1007/s12160-008-9042-y] [Medline: 18633685]
21. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
22. 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]
23. Matheson GO, Klügl M, Engebretsen L, Bendiksen F, Blair SN, Börjesson M, et al. Prevention and management of noncommunicable disease: the IOC Consensus Statement, Lausanne 2013. Clin J Sport Med 2013 Nov;23(6):419-429. [doi: 10.1097/JSM.0000000000000038] [Medline: 24169298]
24. Kelley T, Kelley D. Creative Confidence: Unleashing the Creative Potential Within All of Us. New York: Crown Business; 2013.
25. Stanford d.school: Design Thinking Bootcamp Bootleg. 2011. URL: http://dschool.stanford.edu/wp-content/uploads/2011/03/BootcampBootleg2010v2SLIM.pdf [accessed 2016-11-22] [WebCite Cache ID 6mIYu233k]
26. Norman D, Draper S. User-centred system design: new perspective on human-computer interaction. Hillsdale, NJ: Lawrence Earlbaum Associate; 1986.
27. Hermawati S, Lawson G. Managing obesity through mobile phone applications: a state-of-the-art review from a user-centred design perspective. Pers Ubiquit Comput 2014;18(8):2003-2023.
28. Martin R. Design of Business: Why Design Thinking is the Next Competitive Advantage. Boston: Harvard Business Press; 2009.
29. Dow SP, Glassco A, Kass J, Schwarz M, Schwartz DL, Klemmer SR. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. ACM Trans Comput-Hum Interact 2010 Dec 01;17(4):1-24. [doi: 10.1145/1879831.1879836]
30. Katoppo M, Sudradjat I. Combining Participatory Action Research (PAR) and Design Thinking (DT) as an alternative research method in architecture. Procedia Soc Behav Sci 2015 May;184:118-125. [doi: 10.1016/j.sbspro.2015.05.069]
31. Quade S, Schlueter O. Adapting design thinking for media prototyping: innovative collaboration at universities and workplaces. In: ICERI2014 Proceedings. 2014 Nov 17 Presented at: 7th International Conference of Education, Research and Innovation; Nov 17-19, 2014; Seville, Spain p. 5221-5229.
32. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
33. Coughlin SS, Whitehead M, Sheats JQ, Mastromonico J, Hardy D, Smith SA. Smartphone applications for promoting healthy diet and nutrition: a literature review. Jacobs J Food Nutr 2015;2(3):21 [FREE Full text] [Medline: 26819969]
34. Horvath KJ, Ecklund AM, Hunt SL, Nelson TF, Toomey TL. Developing Internet-based health interventions: a guide for public health researchers and practitioners. J Med Internet Res 2015;17(1):e28 [FREE Full text] [doi: 10.2196/jmir.3770] [Medline: 25650702]
35. van Gemert-Pijnen JE, 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;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
36. Van Velsen L, Wentzel J, Van Gemert-Pijnen JE. Designing eHealth that matters via a multidisciplinary requirements development approach. JMIR Res Protoc 2013;2(1):e21 [FREE Full text] [doi: 10.2196/resprot.2547] [Medline: 23796508]
37. Ludden GD, van Rompay TJ, Kelders SM, van Gemert-Pijnen JE. How to increase reach and adherence of web-based interventions: a design research viewpoint. J Med Internet Res 2015;17(7):e172 [FREE Full text] [doi: 10.2196/jmir.4201] [Medline: 26163456]
38. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
39. Fjeldsoe BS, Miller YD, O'Brien JL, Marshall AL. Iterative development of MobileMums: a physical activity intervention for women with young children. Int J Behav Nutr Phys Act 2012;9:151 [FREE Full text] [doi: 10.1186/1479-5868-9-151] [Medline: 23256730]
40. Waterlander W, Whittaker R, McRobbie H, Dorey E, Ball K, Maddison R, et al. Development of an evidence-based mHealth weight management program using a formative research process. JMIR Mhealth Uhealth 2014;2(3):e18 [FREE Full text] [doi: 10.2196/mhealth.2850] [Medline: 25098337]

41. Whittaker R, Dorey E, Bramley D, Bullen C, Denny S, Elley CR, et al. A theory-based video messaging mobile phone intervention for smoking cessation: randomized controlled trial. J Med Internet Res 2011;13(1):e10 [FREE Full text] [doi: 10.2196/jmir.1553] [Medline: 21371991]
42. Whittaker R, Merry S, Stasiak K, McDowell H, Doherty I, Shepherd M, et al. MEMO--a mobile phone depression prevention intervention for adolescents: development process and postprogram findings on acceptability from a randomized controlled trial. J Med Internet Res 2012;14(1):e13 [FREE Full text] [doi: 10.2196/jmir.1857] [Medline: 22278284]
43. Mummah SA, Mathur M, King AC, Gardner CD, Sutton S. Mobile technology for vegetable consumption: a randomized controlled pilot study in overweight adults. JMIR Mhealth Uhealth 2016 May;4(2):e51 [FREE Full text] [doi: 10.2196/mhealth.5146] [Medline: 27193036]
44. Brown T, Wyatt J. Design thinking for social innovation. Stanford Soc Innov Rev 2010;30(5):29-53.
45. Thondoo M, Strachan DL, Nakirunda M, Ndima S, Muiambo A, Källander K, et al. Potential roles of Mhealth for community health workers: formative research with end users in Uganda and Mozambique. JMIR Mhealth Uhealth 2015;3(3):e76 [FREE Full text] [doi: 10.2196/mhealth.4208] [Medline: 26206419]
46. Hammersley M. Reading Ethnographic Research: A Critical Guide. London: Longman; 1990.
47. Peeters M, Megens C, Hummels C, Brombacher A. Experiential probes: probing for emerging behavior patterns in everyday life. 2013 Aug 26 Presented at: 2013 IASDR Conference: Consilience and Innovation in Design; Aug 26-30, 2013; Tokyo, Japan.
48. Gasparini A. Perspective and use of empathy in design thinking. 2015 Presented at: The Eighth International Conference on Advances in Computer-Human Interactions; Feb 22-27, 2015; Lisbon, Portugal.
49. Lynna GS, Akgunb AE. Project visioning: its components and impact on new product success. J Prod Innov Manag 2001 Nov;18(6):374-387. [doi: 10.1111/1540-5885.1860374]
50. Duhigg C. The power of habit: why we do what we do in life and business. New York: Random House Trade Paperbacks; 2012.
51. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
52. Bandura A. Social Foundations of Thought and Action. Englewood Cliffs, NJ: Prentice-Hall; 1986.
53. Robinson T. Stealth interventions for obesity prevention and control: motvating behavior change. In: Dubé L, Bechara A, Dagher A, Drewnowski A, Lebel J, James P, editors. Obesity Prevention: The Role of Brain and Society on Individual Behavior. New York: Elsevier Inc; 2010:319-327.
54. Wansink B. Mindless Eating: Why We Eat More Than We Think. New York: Bantam Books; 2006.
55. Ariely D. Predictably Irrational: The Hidden Forces that Shape our Decisions. New York: HarperCollins; 2008.
56. Bandura A. Self-Efficacy: The Exercise of Control. New York: Freeman and Company; 1997.
57. Lepper M, Master A, Yow W. Intrinsic motivation in education. In: Maehr ML, Karabenick SA, Urdan TC, editors. Advances in Motivation and Achievement, Volume 15: Social Psychological Perspectives. Bingley: Emerald; 2008:521-555.
58. Robinson TN, Borzekowski DLG. Effects of the SMART classroom curriculum to reduce child and family screen time. J Communication 2006 Mar;56(1):1-26. [doi: 10.1111/j.1460-2466.2006.00001.x]
59. Robinson TN. Save the world, prevent obesity: piggybacking on existing social and ideological movements. Obesity (Silver Spring) 2010 Feb;18 Suppl 1:S17-S22 [FREE Full text] [doi: 10.1038/oby.2009.427] [Medline: 20107456]
60. Dweck C. Mindset: The New Psychology of Success. New York: Ballantine Books; 2008.
61. Crum AJ, Langer EJ. Mind-set matters: exercise and the placebo effect. Psychol Sci 2007 Feb;18(2):165-171. [doi: 10.1111/j.1467-9280.2007.01867.x] [Medline: 17425538]
62. Petrie K, Cameron L, Ellis C, Buick D, Weinman J. Changing illness perceptions after myocardial infarction: an early intervention randomized controlled trial. Psychosom Med 2002;64(4):580-586.
63. Kelley T, Littman J. The Ten Faces of Innovation: IDEO's Strategies for Defeating the Devil's Advocate and Driving Creativity Throughout Your Organization. New York: Random House; 2005.
64. Pelled LH, Eisenhardt KM, Xin KR. Exploring the black box: an analysis of work group diversity, conflict, and performance. Admin Sci Quart 1999 Mar;44(1):1. [doi: 10.2307/2667029]
65. Dow S, Heddleston K, Klemmer S. The efficacy of prototyping under time constraints. In: Proceedings of the Seventh ACM Conference on Creativity and Cognition. 1640260: ACM; 2009 Presented at: C&C '09: Seventh ACM Conference on Creativity and Cognition; Oct 27-30, 2009; Berkeley, CA p. 165-174.
66. Vredenburg K, Isensee S, Righi C. User-Centered Design: An Integrated Approach. Upper Saddle River, NJ: Prentice Hall; 2002.
67. Schrage M. Serious Play: How the World's Best Companies Simulate to Innovate. Brighton: Harvard Business School Press; 1999.
68. Seidel V, Fixon S. Adopting design thinking in novice mutlidisciplinary teams: the application and limits of design methods and reflexive practices. J Prod Innov Manage 2013;30:19-33.
69. Spruijt-Metz D, Nilsen W. Dynamic models of behavior for just-in-time adaptive interventions. Lect Notes Comput Sc 2014;13(3):13-17.

70. Nahum-Shani S, Smith S, Tewari A, Witkiewitz K, Collins L, Spring B. Just-in-Time Adaptive Interventions (JITAIs): An Organizing Framework for Ongoing Health Behavior Support (Technical Report No. 14-126). University Park, PA: The Methodology Center; 2014:1-37.
71. Ries E. The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses. New York: Crown Business; 2011.
72. King AC, Hekler EB, Grieco LA, Winter SJ, Sheats JL, Buman MP, et al. Harnessing different motivational frames via mobile phones to promote daily physical activity and reduce sedentary behavior in aging adults. PLoS One 2013;8(4):e62613 [FREE Full text] [doi: 10.1371/journal.pone.0062613] [Medline: 23638127]
73. Eldridge SM, Ashby D, Feder GS, Rudnicka AR, Ukoumunne OC. Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clin Trials 2004 Feb 01;1(1):80-90. [doi: 10.1191/1740774504cn006rr]
74. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
75. Nollen NL, Mayo MS, Carlson SE, Rapoff MA, Goggin KJ, Ellerbeck EF. Mobile technology for obesity prevention: a randomized pilot study in racial- and ethnic-minority girls. Am J Prev Med 2014 Apr;46(4):404-408 [FREE Full text] [doi: 10.1016/j.amepre.2013.12.011] [Medline: 24650843]
76. Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. J Clin Nurs 2003 Jan;12(1):77-84. [Medline: 12519253]
77. Kraemer HC, Frank E, Kupfer DJ. Moderators of treatment outcomes: clinical, research, and policy importance. JAMA 2006 Sep 13;296(10):1286-1289. [doi: 10.1001/jama.296.10.1286] [Medline: 16968853]
78. Schulz KF. Randomized controlled trials. Clin Obstet Gynecol 1998 Jun;41(2):245-256. [Medline: 9646957]
79. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995 Feb 1;273(5):408-412. [Medline: 7823387]
80. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. JAMA 1994 Jul 13;272(2):125-128. [Medline: 8015122]
81. Herman A, Botser IB, Tenenbaum S, Chechick A. Intention-to-treat analysis and accounting for missing data in orthopaedic randomized clinical trials. J Bone Joint Surg Am 2009 Sep;91(9):2137-2143. [doi: 10.2106/JBJS.H.01481] [Medline: 19723990]
82. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999 Sep 11;319(7211):670-674. [doi: 10.1136/bmj.319.7211.670]
83. Moher D, Hopewell S, Schulz K, Montori V, Gøtzsche P, Devereaux P. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol 2010;63(8):e1-e37. [doi: 10.1016/j.jclinepi.2010.03.004]
84. Schneider F, Schulz DN, Pouwels LH, de Vries H, van Osch LA. The use of a proactive dissemination strategy to optimize reach of an internet-delivered computer tailored lifestyle intervention. BMC Public Health 2013;13:721 [FREE Full text] [doi: 10.1186/1471-2458-13-721] [Medline: 23914991]
85. Gersick CJ. Marking time: predictable transitions in task groups. Acad Manage J 1989 Jun 01;32(2):274-309. [doi: 10.2307/256363]
86. Kelders SM, Van Gemert-Pijnen JE, Werkman A, Nijland N, Seydel ER. Effectiveness of a Web-based intervention aimed at healthy dietary and physical activity behavior: a randomized controlled trial about users and usage. J Med Internet Res 2011;13(2):e32 [FREE Full text] [doi: 10.2196/jmir.1624] [Medline: 21493191]
87. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]

Abbreviations

d.school: Stanford University Hasso Plattner Institute of Design
eHealth: electronic health
IDEAS: Integrate, Design, Assess, and Share
JITAI: just-in-time, adaptive intervention
mHealth: mobile health
MVP: minimum viable product
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 03.05.16; peer-reviewed by G Munro, A Atienza, M Peeples; comments to author 16.06.16; revised version received 25.09.16; accepted 12.10.16; published 16.12.16.

Please cite as:
Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S
IDEAS (Integrate, Design, Assess, and Share): A Framework and Toolkit of Strategies for the Development of More Effective Digital Interventions to Change Health Behavior
J Med Internet Res 2016;18(12):e317
URL: http://www.jmir.org/2016/12/e317/
doi:10.2196/jmir.5927
PMID:27986647

Original Paper

Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View

Wei Luo[1*], PhD; Dinh Phung[2*], PhD; Truyen Tran[2*], PhD; Sunil Gupta[2*], PhD; Santu Rana[2*], PhD; Chandan Karmakar[2*], PhD; Alistair Shilton[2*], PhD; John Yearwood[2*], PhD; Nevenka Dimitrova[3*], PhD; Tu Bao Ho[4*], PhD; Svetha Venkatesh[2*], PhD; Michael Berk[2*], PhD, FRACP

[1]Centre for Pattern Recognition and Data Analytics, School of Information Technology, Deakin University, Geelong, Australia
[2]Deakin University, Geelong, Australia
[3]Philips Research, Briarcliff Manor, NY, United States
[4]Japan Advanced Institute of Science and Technology, Nomi, Japan
[*]all authors contributed equally

Corresponding Author:
Wei Luo, PhD
Centre for Pattern Recognition and Data Analytics
School of Information Technology
Deakin University
Building KA
75 Pigdons Road
Geelong, 3220
Australia
Phone: 61 3 5227 3096
Fax: 61 3 5227 3096
Email: wei.luo@deakin.edu.au

Abstract

Background: As more and more researchers are turning to big data for new opportunities of biomedical discoveries, machine learning models, as the backbone of big data analysis, are mentioned more often in biomedical journals. However, owing to the inherent complexity of machine learning methods, they are prone to misuse. Because of the flexibility in specifying machine learning models, the results are often insufficiently reported in research articles, hindering reliable assessment of model validity and consistent interpretation of model outputs.

Objective: To attain a set of guidelines on the use of machine learning predictive models within clinical settings to make sure the models are correctly applied and sufficiently reported so that true discoveries can be distinguished from random coincidence.

Methods: A multidisciplinary panel of machine learning experts, clinicians, and traditional statisticians were interviewed, using an iterative process in accordance with the Delphi method.

Results: The process produced a set of guidelines that consists of (1) a list of reporting items to be included in a research article and (2) a set of practical sequential steps for developing predictive models.

Conclusions: A set of guidelines was generated to enable correct application of machine learning models and consistent reporting of model specifications and results in biomedical research. We believe that such guidelines will accelerate the adoption of big data analysis, particularly with machine learning methods, in the biomedical research community.

(J Med Internet Res 2016;18(12):e323) doi:10.2196/jmir.5870

KEYWORDS

machine learning; clinical prediction rule; guideline

Introduction

Big data is changing every industry. Medicine is no exception. With rapidly growing volume and diversity of data in health care and biomedical research, traditional statistical methods often are inadequate. By looking into other industries where modern machine learning techniques play central roles in dealing with big data, many health and biomedical researchers have started applying machine learning to extract valuable insights from ever-growing biomedical databases, in particular with predictive models [1,2]. The flexibility and prowess of machine learning models also enable us to leverage novel but extremely valuable sources of information, such as wearable device data and electronic health record data [3].

Despite its popularity, it is difficult to find a universally agreed-upon definition for machine learning. Arguably, many machine learning methods can be traced back as far as 30 years ago [4]. However, machine learning started making a broad impact only in the last 10 years. The reviews by Jordan and Mitchell [5] and Ghahramani [6] provide accessible overviews for machine learning. In this paper, we focus on machine learning predictive methods and models. These include random forest, support vector machines, and other methods listed in Multimedia Appendix 1. They all share an important difference from the traditional statistical methods such as logistic regression or analysis of variance—the ability to make accurate predictions on unseen data. To optimize the prediction accuracy, often the methods do not attempt to produce interpretable models. This also allows them to handle a large number of variables common in most big data problems.

Accompanying the flexibility of emerging machine learning techniques, however, is uncertainty and inconsistency in the use of such techniques. Machine learning, owing to its intrinsic mathematical and algorithmic complexity, is often considered a “black magic” that requires a delicate balance of a large number of conflicting factors. This, together with inadequate reporting of data sources and modeling process, makes research results reported in many biomedical papers difficult to interpret. It is not rare to see potentially spurious conclusions drawn from methodologically inadequate studies [7-11], which in turn compromises the credibility of other valid studies and discourages many researchers who could benefit from adopting machine learning techniques.

Most pitfalls of applying machine learning techniques in biomedical research originate from a small number of common issues, including data leakage [12] and overfitting [13-15], which can be avoided by adopting a set of best practice standards. Recognizing the urgent need for such a standard, we created a minimum list of reporting items and a set of guidelines for optimal use of predictive models in biomedical research.

Methods

Panel of Experts

In 2015, a multidisciplinary panel was assembled to cover expertise in machine learning, traditional statistics, and biomedical applications of these methods. The candidate list was generated in two stages. The panel grew from a number of active machine learning researchers attending international conferences including the Asian Conference on Machine Learning, the Pacific Asia Conference on Knowledge Discovery and Data Mining, and the International Conference on Pattern Recognition. The responders were then asked to nominate additional researchers who apply machine learning in biomedical research. Effort was exercised to include researchers from different continents. Researchers from the list were approached through emails for joining the panel and/or recommending colleagues to be included. Two declined the invitation.

The final panel included 11 researchers from 3 institutions on 3 different continents. Each panelist had experience and expertise in machine learning projects in biomedical applications and has learned from common pitfalls. The areas of research expertise included machine learning, data mining, computational intelligence, signal processing, information management, bioinformatics, and psychiatry. On average, each panel member had 8.5 years’ experience in either developing or applying machine learning methods. The diversity of the panel was reflected by the members’ affiliation with 3 different institutions across 3 continents.

Development of Guidelines

Using an iterative process, the panel produced a set of guidelines that consists of (1) a list of reporting items to be included in a research article and (2) a set of practical sequential steps for developing predictive models. The Delphi method was used to generate the list of reporting items.

The panelists were interviewed with multiple iterations of emails. Email 1 asked panelists to list topics to be covered in the guidelines. An aggregated topic list was generated. Email 2 asked each panelist to review the scope of the list and state his or her recommendation for each topic in the aggregated list. Later iterations of email interviews were organized to evolve the list until all experts agreed on the list. Because of the logistic complexity of coordinating the large panel, we took a grow-shrink approach. In the growing phase, all suggested items were included, even an item suggested by only 1 panelist. In the shrinking phase, any item opposed by a panelist was excluded. As it turned out, most items were initially suggested by a panelist but seconded by other panelists, suggesting the importance of the group effort for covering most important topics.

The practical steps were developed by machine learning experts in their respective areas and finally approved by the panel. During the process, the panelists consulted extensively the broad literature on machine learning and predictive model in particular [16-18].

Results

A total of 4 iterations of emails resulted in the final form of the guidelines. Email 1 generated diverse responses in terms of topics. However the final scope was generally agreed upon. For email 2, most panelists commented on only a subset of topics (mostly the ones suggested by themselves). No recommendations generated significant disagreement except for minor wording decisions and quantifying conditions.

The final results included a list of reporting items (Tables 1-5,Textboxes 1-4, and Figure 1) and a template flowchart for reporting data used for training and testing predictive models, including both internal validation and external validation (Figure 2).

Recognizing the broad meaning of the term “machine learning,” we distinguish essential items from desirable items (using appropriate footnotes in the tables). The essential items should be included in any report, unless there is a strong reason indicating otherwise; the desirable items should be reported whenever applicable.

Figure 1. Steps to identify the prediction problem.

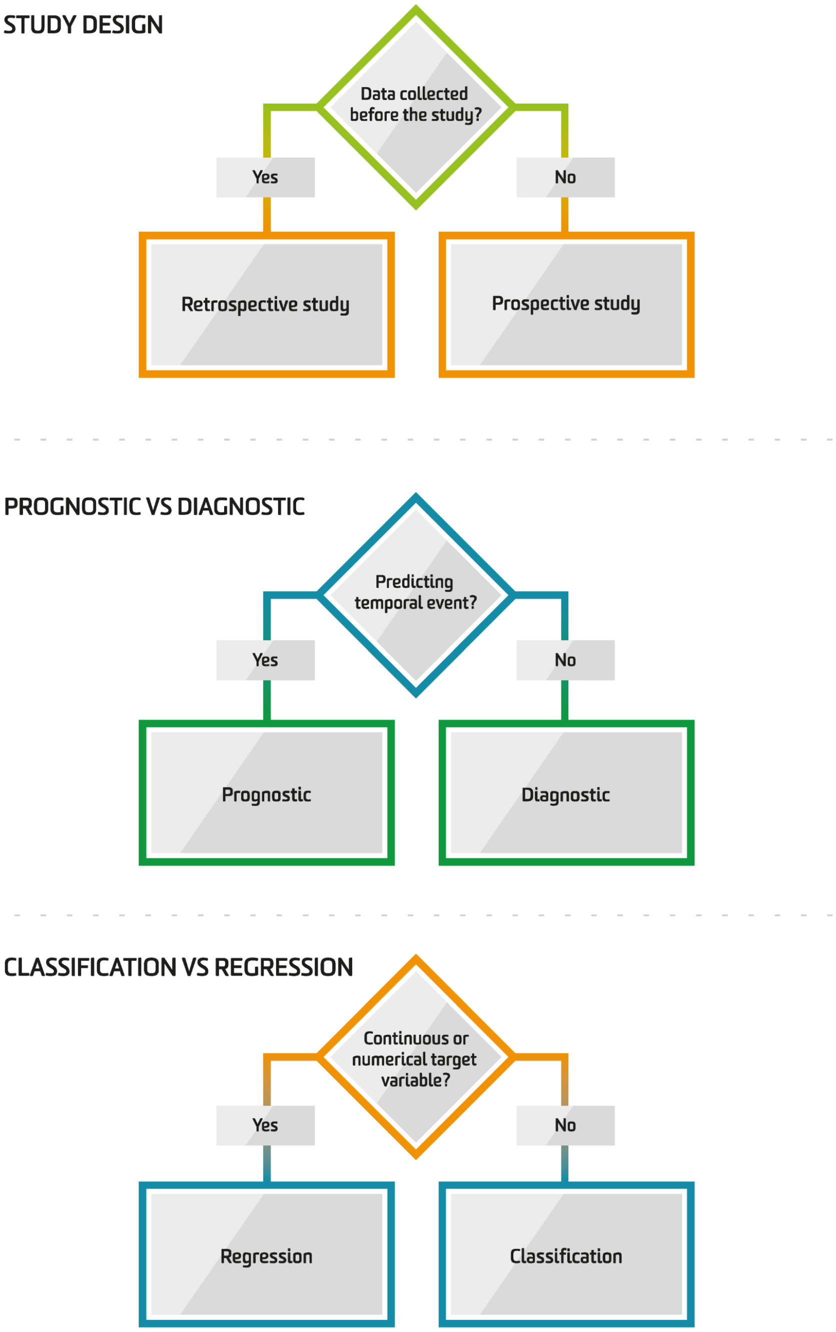


Table 1. Items to include when reporting predictive models in biomedical research: title and abstract.

Item number	Section	Topic	Checklist item
1	Title	Nature of study	Identify the report as introducing a predictive model
2	Abstract	Structured summary	Background Objectives Data sources Performance metrics of the predictive model or models, in both point estimates and confidence intervals Conclusion including the practical value of the developed predictive model or models

Table 2. Items to include when reporting predictive models in biomedical research: introduction section.

Item number	Topic	Checklist item
3	Rationale	Identify the clinical goal Review the current practice and prediction accuracy of any existing models
4	Objectives	State the nature of study being predictive modeling, defining the target of prediction Identify how the prediction problem may benefit the clinical goal

Figure 2. Information flow in the predictive modelling process.

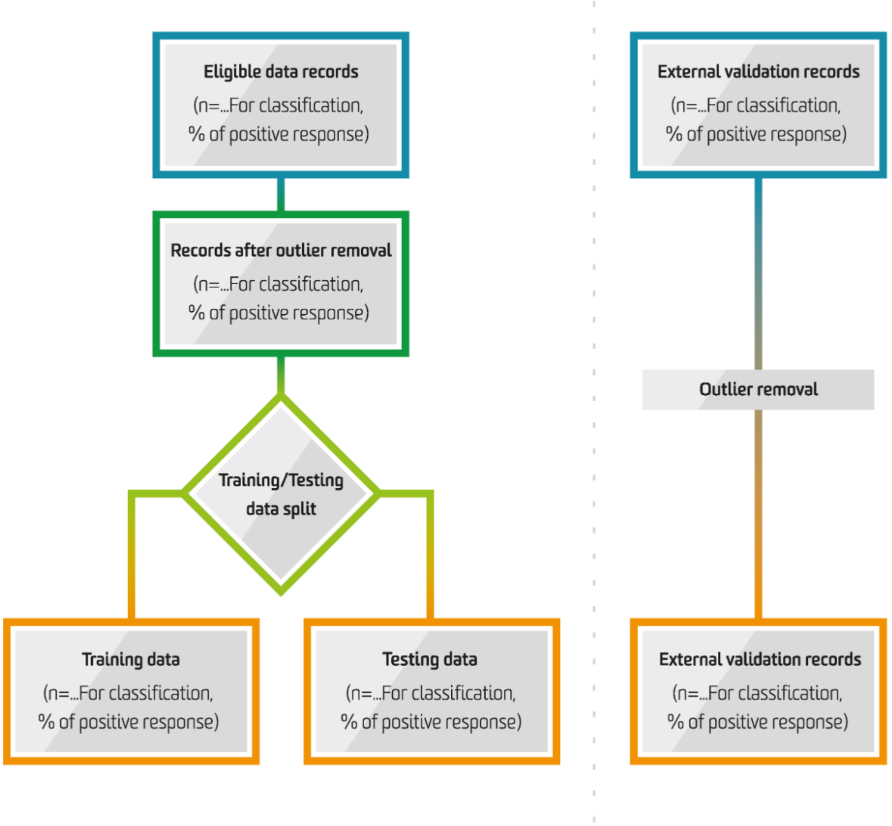


Table 3. Items to include when reporting predictive models in biomedical research: methods section.

Item number	Topic	Checklist item
5	Describe the setting	Identify the clinical setting for the target predictive model.
		Identify the modeling context in terms of facility type, size, volume, and duration of available data.
6	Define the prediction problem	Define a measurement for the prediction goal (per patient or per hospitalization or per type of outcome).
		Determine that the study is retrospective or prospective.[a]
		Identify the problem to be prognostic or diagnostic.
		Determine the form of the prediction model: (1) classification if the target variable is categorical, (2) regression if the target variable is continuous, (3) survival prediction if the target variable is the time to an event.
		Translate survival prediction into a regression problem, with the target measured over a temporal window following the time of prediction.
		Explain practical costs of prediction errors (eg, implications of underdiagnosis or overdiagnosis).
		Defining quality metrics for prediction models.[b]
		Define the success criteria for prediction (eg, based on metrics in internal validation or external validation in the context of the clinical problem).
7	Prepare data for model building	Identify relevant data sources and quote the ethics approval number for data access.
		State the inclusion and exclusion criteria for data.
		Describe the time span of data and the sample or cohort size.
		Define the observational units on which the response variable and predictor variables are defined.
		Define the predictor variables. Extra caution is needed to prevent information leakage from the response variable to predictor variables.[c]
		Describe the data preprocessing performed, including data cleaning and transformation. Remove outliers with impossible or extreme responses; state any criteria used for outlier removal.
		State how missing values were handled.
		Describe the basic statistics of the dataset, particularly of the response variable. These include the ratio of positive to negative classes for a classification problem and the distribution of the response variable for regression problem.
		Define the model validation strategies. Internal validation is the minimum requirement; external validation should also be performed whenever possible.
		Specify the internal validation strategy. Common methods include random split, time-based split, and patient-based split.
		Define the validation metrics. For regression problems, the normalized root-mean-square error should be used. For classification problems, the metrics should include sensitivity, specificity, positive predictive value, negative predictive value, area under the ROC[d] curve, and calibration plot [19].[e]
		For retrospective studies, split the data into a derivation set and a validation set. For prospective studies, define the starting time for validation data collection.
8	Build the predictive model	Identify independent variables that predominantly take a single value (eg, being zero 99% of the time).
		Identify and remove redundant independent variables.
		Identify the independent variables that may suffer from the perfect separation problem.[f]
		Report the number of independent variables, the number of positive examples, and the number of negative examples.
		Assess whether sufficient data are available for a good fit of the model. In particular, for classification, there should be a sufficient number of observations in both positive and negative classes.
		Determine a set of candidate modeling techniques (eg, logistic regression, random forest, or deep learning). If only one type of model was used, justify the decision for using that model.[g]
		Define the performance metrics to select the best model.
		Specify the model selection strategy. Common methods include K-fold validation or bootstrap to estimate the lost function on a grid of candidate parameter values. For K-fold validation, proper stratification by the response variable is needed.[h]
		For model selection, include discussion on (1) balance between model accuracy and model simplicity or interpretability, and (2) the familiarity with the modeling techniques of the end user.[i]

[a] See Figure 1.

[b] See some examples in Multimedia Appendix 2.

[c] See Textbox 1.

[d] ROC: receiver operating characteristic.

[e] Also see Textbox 2.

[f] See Textbox 3.

[g] See Multimedia Appendix 1 for some common methods and their strengths and limitations.

[h] See Textbox 4.

[i] A desirable but not mandatory item.

Table 4. Items to include when reporting predictive models in biomedical research: results section.

Item number	Topic	Checklist item
9	Report the final model and performance	Report the predictive performance of the final model in terms of the validation metrics specified in the methods section. If possible, report the parameter estimates in the model and their confidence intervals. When the direct calculation of confidence intervals is not possible, report nonparametric estimates from bootstrap samples. Comparison with other models in the literature should be based on confidence intervals. Interpretation of the final model. If possible, report what variables were shown to be predictive of the response variable. State which subpopulation has the best prediction and which subpopulation is most difficult to predict.

Table 5. Items to include when reporting predictive models in biomedical research: discussion section.

Item number	Topic	Checklist item
10	Clinical implications	Report the clinical implications derived from the obtained predictive performance. For example, report the dollar amount that could be saved with better prediction. How many patients could benefit from a care model leveraging the model prediction? And to what extent?
11	Limitations of the model	Discuss the following potential limitations: • Assumed input and output data format • Potential pitfalls in interpreting the model[a] • Potential bias of the data used in modeling • Generalizability of the data
12	Unexpected results during the experiments	Report unexpected signs of coefficients, indicating collinearity or complex interaction between predictor variables[a]

[a] Desirable but not mandatory items.

Textbox 1. Data leakage problem.

Leakage refers to the unintended use of known information as unknown. There are two kinds of leakage: outcome leakage and validation leakage. In outcome leakage, independent variables incorporate elements that can be used to easily infer outcomes. For example, a risk factor that spans into the future may be used to predict the future itself. In the validation leakage, ground truth from the training set may propagate to the validation set. For example, when the same patient is used in both training and validation, the future outcome in the training may overlap with the future outcome in the validation. In both leakage cases, the performance obtained is overoptimistic.

Textbox 2. Calibration.

Calibration of a prediction model refers to the agreement between the predictions made by the model and the observed outcomes. As an example, if the prediction model predicts 70% risk of mortality in the next 1 year for a patient with lung cancer, then the model is well calibrated if in our dataset approximately 70% of patients with lung cancer die within the next 1 year.

Often, the regularized prediction may create bias in a model. Therefore, it is advisable to check for the calibration. In the case of regression models, the calibration can be easily assessed graphically by marking prediction scores on the x-axis and the true outcomes on the y-axis. In the case of binary classification, the y-axis has only 0 and 1 values; however, smoothing techniques such as LOESS algorithm may be used to estimate the observed probabilities for the outcomes. In a more systematic way, one can perform the Hosmer-Lemeshow test to measure the goodness of fit of the model. The test assesses whether the observed event rates match the predicted event rates in subgroups of the model population.

Textbox 3. Perfect separation problem.

When a categorical predictor variable can take an uncommon value, there may be only a small number of observations having that value. In a classification problem, these few observations by chance may have the same response value. Such "perfect" predictors may cause overfitting, especially when tree-based models are used. Therefore, special treatment is required.

One conservative approach is to remove all dummy variables corresponding to rare categories. We recommend a cutoff of 10 observations.

For modeling methods with feature selection built in, an alternative approach is to first fit a model with all independent variables. If the resulting model is only marginally influenced by the rare categories, then the model can be kept. Otherwise, the rare categories showing high "importance" score are removed and the model refitted.

Textbox 4. K-fold cross-validation.

K-fold validation refers to the practice of splitting the derivation data into K equal parts. The model is then trained on K−1 parts and validated on the remaining part. The process is repeated K times. The average results for K-folds are then reported. For small classes and rare categorical factors, stratified K-fold splitting should be used to ensure the equal presence of these classes and factors in each fold.

Discussion

We have generated a set of guidelines that will enable correct application of machine learning models and consistent reporting of model specifications and results in biomedical research.

Because of the broad range of machine learning methods that can be used in biomedical applications, we involved a large number of stakeholders, as either developers of machine learning methods or users of these methods in biomedicine research.

The guidelines here cover most popular machine learning methods appearing in biomedical studies. We believe that such guidelines will accelerate the adoption of big data analysis, particularly with machine learning methods, in the biomedical research community.

Although the proposed guidelines result from a voluntary effort without dedicated funding support, we still managed to assemble a panel of researchers from multiple disciplines, multiple institutions, and multiple continents. We hope the guidelines can result in more people contributing their knowledge and experience in the discussion.

As machine learning is a rapidly developing research area, the guidelines are not expected to cover every aspect of the modeling process. The guidelines are expected to evolve as research in biomedicine and machine learning progresses.

Acknowledgments

This project is partially supported by the Telstra-Deakin COE in Big data and Machine Learning.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Guidelines for fitting some common predictive models.

[PDF File (Adobe PDF File), 77KB - jmir_v18i12e323_app1.pdf]

Multimedia Appendix 2

Glossary.

[PDF File (Adobe PDF File), 21KB - jmir_v18i12e323_app2.pdf]

References

1. Ayaru L, Ypsilantis P, Nanapragasam A, Choi RC, Thillanathan A, Min-Ho L, et al. Prediction of outcome in acute lower gastrointestinal bleeding using gradient boosting. PLoS One 2015;10(7):e0132485 [FREE Full text] [doi: 10.1371/journal.pone.0132485] [Medline: 26172121]
2. Ogutu J, Schulz-Streeck T, Piepho HP. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. BMC Proc 2012;6(Suppl 2):S10. [doi: 10.1186/1753-6561-6-S2-S10]
3. Tran T, Luo W, Phung D, Harvey R, Berk M, Kennedy RL, et al. Risk stratification using data from electronic medical records better predicts suicide risks than clinician assessments. BMC Psychiatry 2014 Mar 14;14:76 [FREE Full text] [doi: 10.1186/1471-244X-14-76] [Medline: 24628849]
4. Breiman L, Friedman J, Stone C, Olshen R. Classification and regression trees. New York: Chapman & Hall; 1984.

5. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science 2015 Jul 17;349(6245):255-260. [doi: 10.1126/science.aaa8415] [Medline: 26185243]
6. Ghahramani Z. Probabilistic machine learning and artificial intelligence. Nature 2015 May 28;521(7553):452-459. [doi: 10.1038/nature14541] [Medline: 26017444]
7. Bone D, Goodwin MS, Black MP, Lee C, Audhkhasi K, Narayanan S. Applying machine learning to facilitate autism diagnostics: pitfalls and promises. J Autism Dev Disord 2015 May;45(5):1121-1136 [FREE Full text] [doi: 10.1007/s10803-014-2268-6] [Medline: 25294649]
8. Metaxas P, Mustafaraj E, Gayo-Avello D. How (not) to predict elections. 2011 Presented at: 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing; 2011; Boston.
9. Jungherr A, Jurgens P, Schoen H. Why the pirate party won the German election of 2009 or the trouble with predictions: a response to Tumasjan A, Sprenger TO, Sander PG, & Welpe IM “Predicting elections with Twitter: what 140 characters reveal about political sentiment”. Social Science Computer Review 2011 Apr 25;30(2):229-234. [doi: 10.1177/0894439311404119]
10. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
11. Foster KR, Koprowski R, Skufca JD. Machine learning, medical diagnosis, and biomedical engineering research - commentary. Biomed Eng Online 2014 Jul 05;13:94 [FREE Full text] [doi: 10.1186/1475-925X-13-94] [Medline: 24998888]
12. Smialowski P, Frishman D, Kramer S. Pitfalls of supervised feature selection. Bioinformatics 2010 Feb 1;26(3):440-443 [FREE Full text] [doi: 10.1093/bioinformatics/btp621] [Medline: 19880370]
13. Babyak M. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 2004;66(3):411-421. [Medline: 15184705]
14. Hawkins DM. The problem of overfitting. J Chem Inf Comput Sci 2004;44(1):1-12. [doi: 10.1021/ci0342472] [Medline: 14741005]
15. Subramanian J, Simon R. Overfitting in prediction models - is it a problem only in high dimensions? Contemp Clin Trials 2013 Nov;36(2):636-641. [doi: 10.1016/j.cct.2013.06.011] [Medline: 23811117]
16. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference and prediction. New York, NY: Springer; 2009.
17. Kuhn M, Johnson K. Applied predictive modeling. Berlin: Springer; 2013.
18. Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating (Statistics for Biology and Health). New York: Springer; 2009.
19. Cox D. Two further applications of a model for binary regression. Biometrika 1958 Dec;45(3/4):562-565. [doi: 10.2307/2333203]
20. Quinlan J. Simplifying decision trees. Int J Man Mach Stud 1987 Sep;27(3):221-234. [doi: 10.1016/S0020-7373(87)80053-6]
21. Quinlan J. Induction of decision trees. Mach Learn 1986 Mar;1(1):81-106. [doi: 10.1007/BF00116251]
22. Podgorelec V, Kokol P, Stiglic B, Rozman I. Decision trees: an overview and their use in medicine. J Med Syst 2002 Oct;26(5):445-463. [Medline: 12182209]
23. Kotsiantis S. Decision trees: a recent overview. Artif Intell Rev 2011 Jun 29;39(4):261-283. [doi: 10.1007/s10462-011-9272-4]
24. Kingsford C, Salzberg SL. What are decision trees? Nat Biotechnol 2008 Sep;26(9):1011-1013 [FREE Full text] [doi: 10.1038/nbt0908-1011] [Medline: 18779814]
25. Luo W, Gallagher M. Unsupervised DRG upcoding detection in healthcare databases. 2010 Presented at: IEEE International Conference on Data Mining Workshops (ICDMW); 2010; Sydney p. 600-605. [doi: 10.1109/ICDMW.2010.108]
26. Siddique J, Ruhnke GW, Flores A, Prochaska MT, Paesch E, Meltzer DO, et al. Applying classification trees to hospital administrative data to identify patients with lower gastrointestinal bleeding. PLoS One 2015;10(9):e0138987 [FREE Full text] [doi: 10.1371/journal.pone.0138987] [Medline: 26406318]
27. Bae S, Lee SA, Lee SH. Prediction by data mining, of suicide attempts in Korean adolescents: a national study. Neuropsychiatr Dis Treat 2015;11:2367-2375 [FREE Full text] [doi: 10.2147/NDT.S91111] [Medline: 26396521]
28. Satomi J, Ghaibeh AA, Moriguchi H, Nagahiro S. Predictability of the future development of aggressive behavior of cranial dural arteriovenous fistulas based on decision tree analysis. J Neurosurg 2015 Jul;123(1):86-90. [doi: 10.3171/2014.10.JNS141429] [Medline: 25859811]
29. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
30. Boulesteix A, Janitza S, Kruppa J, König I. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. WIREs Data Mining Knowl Discov 2012 Oct 18;2(6):493-507. [doi: 10.1002/widm.1072]
31. Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods 2009 Dec;14(4):323-348 [FREE Full text] [doi: 10.1037/a0016973] [Medline: 19968396]

32. Touw W, Bayjanov JR, Overmars L, Backus L, Boekhorst J, Wels M, et al. Data mining in the Life Sciences with random forest: a walk in the park or lost in the jungle? Brief Bioinform 2013 May;14(3):315-326 [FREE Full text] [doi: 10.1093/bib/bbs034] [Medline: 22786785]
33. Asaoka R, Iwase A, Tsutsumi T, Saito H, Otani S, Miyata K, et al. Combining multiple HRT parameters using the 'Random Forests' method improves the diagnostic accuracy of glaucoma in emmetropic and highly myopic eyes. Invest Ophthalmol Vis Sci 2014 Apr 17;55(4):2482-2490. [doi: 10.1167/iovs.14-14009] [Medline: 24609628]
34. Yoshida T, Iwase A, Hirasawa H, Murata H, Mayama C, Araie M, et al. Discriminating between glaucoma and normal eyes using optical coherence tomography and the 'Random Forests' classifier. PLoS One 2014;9(8):e106117 [FREE Full text] [doi: 10.1371/journal.pone.0106117] [Medline: 25167053]
35. Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc Series B Stat Methodol 1996:267-288.
36. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. Journal of the Royal Statistical Societyries B (Statistical Methodology) 2011;73(3):273-282. [doi: 10.1111/j.1467-9868.2011.00771.x]
37. Vidaurre D, Bielza C, Larrañaga P. A survey of L1 regression. Int Stat Rev 2013 Oct 24;81(3):361-387. [doi: 10.1111/insr.12023]
38. Hesterberg T, Choi N, Meier L, Fraley C. Least angle and ℓ 1 penalized regression: A review. Statist Surv 2008;2:61-93. [doi: 10.1214/08-SS035]
39. Fujino Y, Murata H, Mayama C, Asaoka R. Applying “lasso” regression to predict future visual field progression in glaucoma patients. Invest Ophthalmol Vis Sci 2015 Apr;56(4):2334-2339. [doi: 10.1167/iovs.15-16445] [Medline: 25698708]
40. Shimizu Y, Yoshimoto J, Toki S, Takamura M, Yoshimura S, Okamoto Y, et al. Toward probabilistic diagnosis and understanding of depression based on functional MRI data analysis with logistic group LASSO. PLoS One 2015;10(5):e0123524 [FREE Full text] [doi: 10.1371/journal.pone.0123524] [Medline: 25932629]
41. Lee T, Chao P, Ting H, Chang L, Huang Y, Wu J, et al. Using multivariate regression model with least absolute shrinkage and selection operator (LASSO) to predict the incidence of Xerostomia after intensity-modulated radiotherapy for head and neck cancer. PLoS One 2014;9(2):e89700 [FREE Full text] [doi: 10.1371/journal.pone.0089700] [Medline: 24586971]
42. Friedman J. Stochastic gradient boosting. Comput Stat Data Anal 2002 Feb;38(4):367-378. [doi: 10.1016/S0167-9473(01)00065-2]
43. Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013;7:21 [FREE Full text] [doi: 10.3389/fnbot.2013.00021] [Medline: 24409142]
44. De'ath G. Boosted trees for ecological modeling and prediction. Ecology 2007 Jan;88(1):243-251. [Medline: 17489472]
45. Mayr A, Binder H, Gefeller O, Schmid M. The evolution of boosting algorithms. From machine learning to statistical modelling. Methods Inf Med 2014;53(6):419-427. [doi: 10.3414/ME13-01-0122] [Medline: 25112367]
46. González-Recio O, Jiménez-Montero JA, Alenda R. The gradient boosting algorithm and random boosting for genome-assisted evaluation in large data sets. J Dairy Sci 2013 Jan;96(1):614-624. [doi: 10.3168/jds.2012-5630] [Medline: 23102953]
47. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press; 2000.
48. Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273-297. [doi: 10.1023/A:1022627411411]
49. Burges C. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 1998;2(2):121-167. [doi: 10.1023/A:1009715923555]
50. Smola A, Schölkopf B. A tutorial on support vector regression. Stat Comput 2004 Aug;14(3):199-222. [doi: 10.1023/B:STCO.0000035301.49549.88]
51. Vapnik V, Mukherjee S. Support Vector Method for Multivariate Density Estimation. 2000 Presented at: Neural Information Processing Systems; 1999; Denver. [doi: 10.1109/ICICIS.2011.41]
52. Manevitz L, Yousef M. One-class SVMs for document classification. J Mach Learn Res 2001;2:139-154.
53. Tsochantaridis I, Hofmann T, Altun Y. Support vector machine learning for interdependent and structured output spaces. 2004 Presented at: the twenty-first international conference on Machine learning; 2004; Banff. [doi: 10.1145/1015330.1015341]
54. Shilton A, Lai DT, Palaniswami M. A division algebraic framework for multidimensional support vector regression. IEEE Trans Syst Man Cybern B Cybern 2010 Apr;40(2):517-528. [doi: 10.1109/TSMCB.2009.2028314] [Medline: 19737676]
55. Shashua A, Levin A. Taxonomy of large margin principle algorithms for ordinal regression problems. Adv Neural Inf Process Syst 2002;15:937-944.

Edited by G Eysenbach; submitted 12.04.16; peer-reviewed by J Gosling, Kfoster; comments to author 27.07.16; revised version received 04.11.16; accepted 23.11.16; published 16.12.16.

Please cite as:
Luo W, Phung D, Tran T, Gupta S, Rana S, Karmakar C, Shilton A, Yearwood J, Dimitrova N, Ho TB, Venkatesh S, Berk M
Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View
J Med Internet Res 2016;18(12):e323
URL: http://www.jmir.org/2016/12/e323/
doi:10.2196/jmir.5870
PMID:27986644

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org