Journal of Medical Internet Research

Impact Factor (2018): 4.945 - ranked #1 medical informatics journal by Impact Factor
Volume 21 (2019), Issue 6 ISSN: 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

Corrigenda and Addendas

XSL•FO
RenderX

Review

Context-Aware Systems for Chronic Disease Patients: Scoping Review

Kathleen Yin[1*], PhD; Liliana Laranjo[1*], MPH, MD, PhD; Huong Ly Tong[1], BHSc; Annie YS Lau[1], PhD; A Baki Kocaballi[1], PhD; Paige Martin[1], BEng (Hons); Sanjyot Vagholkar[2], MBBS (Hons), MPH, PhD, FRACGP; Enrico Coiera[1], MBBS, PhD, FACMI, FACHI

[1]Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

[2]Macquarie University Health Sciences Centre, Macquarie University, Sydney, Australia

[*]these authors contributed equally

Corresponding Author:
Kathleen Yin, PhD
Centre for Health Informatics
Australian Institute of Health Innovation
Macquarie University
75 Talavera Rd
Sydney,
Australia
Phone: 61 2 98502477
Email: kathleen.yin@mq.edu.au

Abstract

Background: Context-aware systems, also known as context-sensitive systems, are computing applications designed to capture, interpret, and use contextual information and provide adaptive services according to the current context of use. Context-aware systems have the potential to support patients with chronic conditions; however, little is known about how such systems have been utilized to facilitate patient work.

Objective: This study aimed to characterize the different tasks and contexts in which context-aware systems for patient work were used as well as to assess any existing evidence about the impact of such systems on health-related process or outcome measures.

Methods: A total of 6 databases (MEDLINE, EMBASE, CINAHL, ACM Digital, Web of Science, and Scopus) were scanned using a predefined search strategy. Studies were included in the review if they focused on patients with chronic conditions, involved the use of a context-aware system to support patients' health-related activities, and reported the evaluation of the systems by the users. Studies were screened by independent reviewers, and a narrative synthesis of included studies was conducted.

Results: The database search retrieved 1478 citations; 6 papers were included, all published from 2009 onwards. The majority of the papers were quasi-experimental and involved pilot and usability testing with a small number of users; there were no randomized controlled trials (RCTs) to evaluate the efficacy of a context-aware system. In the included studies, context was captured using sensors or self-reports, sometimes involving both. Most studies used a combination of sensor technology and mobile apps to deliver personalized feedback. A total of 3 studies examined the impact of interventions on health-related measures, showing positive results.

Conclusions: The use of context-aware systems to support patient work is an emerging area of research. RCTs are needed to evaluate the effectiveness of context-aware systems in improving patient work, self-management practices, and health outcomes in chronic disease patients.

(J Med Internet Res 2019;21(6):e10896) doi:10.2196/10896

KEYWORDS

self-care; medical informatics; mobile applications; chronic disease; self-management

XSL•FO
RenderX

Introduction

Background

As health care moves from the traditional hospital setting to the personal sphere of home and community, individuals are increasingly being encouraged to engage in self-care [1,2]. Sociologists refer to this act of self-care as *patient work*, which involves effort and investment of time on the part of patients or family members to accomplish a health goal [3]. Patient work extends beyond strictly health-related tasks and is shaped by the context of patients' lives and their daily routines [1]. It has been suggested that the use of context-aware technologies may thus better support patient work and improve self-care, as contextual information could trigger more personalized and relevant services or information [4].

Context-aware systems, also known as context-sensitive systems, are computing applications designed to capture, interpret, and use contextual information and provide adaptive services according to the current context of use [4,5]. Context-aware systems may thus harness everything from sensors that capture data indicative of context (such as time, location, and light intensity) to inference mechanisms that interpret and action such data [5]. Even though context-aware systems have been piloted in some health care settings, their impact on health care outcomes remains unclear. Specifically, context-aware systems have mainly been piloted in the hospital setting [4] and for primary prevention [6-8], rarely addressing the context of chronic disease patients' health-related activities in everyday life.

Objectives

The aim of this study was to examine existing literature on interventions using context-aware technologies that support *patient work*. Specifically, we sought to characterize the different tasks and contexts in which such systems were used, as well as assess any existing evidence about their impact on health-related process and outcome measures.

Methods

Search Strategy

A systematic search of the literature was performed in September 2016 and updated in October 2017 on MEDLINE, EMBASE, CINAHL, ACM Digital, Web of Science, and Scopus using search terms regarding patient work, context awareness, and consumer health informatics. The complete search strategy is available in Multimedia Appendix 1. The reference lists of relevant articles were also screened to ensure that all eligible studies were captured. A grey literature search was performed using Google Scholar to capture dissertations, theses, and conference proceedings that met the inclusion criteria.

Study Selection Criteria

In the scope of our study, we focused on context-aware systems that were capable of (1) capturing and processing contextual information (eg, environmental data and user-related features) and (2) using the captured contextual information to provide adaptive services and support patient work tasks in everyday life, either at home or in the community.

Studies were included in the review if they focused on patients with chronic conditions, involved the use of a context-aware system to support patients' health-related activities, and reported the evaluation of the systems by the users.

Studies were excluded if they were not in English or if they focused on health care providers instead of consumers. We also excluded interventions that merely gathered and displayed context information, without using it to adapt system behavior (passive context awareness), as this was outside the scope of this study.

Paper Screening Process

We conducted a 2-phase screening process, initially excluding papers based on their titles and abstracts using a standard screening form, and then rescreening the remaining papers based on the full-text article.

Both phases were conducted by teams of 2 independent reviewers (2 teams in the first phase and 1 in the second). Cohen kappa was used in the full-text paper screening to measure intercoder agreement. Any disagreements in the screening were resolved through discussion and consensus.

Data Extraction and Synthesis Strategy

One reviewer extracted information from the eligible studies into a data extraction form, whereas 2 other reviewers examined the completed form for consistency and accuracy. The following information was collected: first author, year, health domain, study type, participants' characteristics (number, age, and sex ratio), intervention characteristics, health activities (tasks undertaken by patients to achieve health goals), and main findings. To explore how context was utilized by the included studies, we analyzed the related contextual elements based on previous work [4,9]. Context information was grouped into the following dimensions: setting (indoor or outdoor), environmental features (indoor and outdoor attributes, eg, room temperature, humidity, and air pollen), and user features (user-related data captured by the system, eg, physical activity, physiological measurements, and mental state). Time is also considered an important element, but it is often coupled with other dimensions, so it was not analyzed separately in this review. Finally, we characterized the utilization of context for each study (ie, the adaptive services provided by the system based on contextual data).

Our study design follows the guidelines for a scoping review proposed by Arksey and O'Malley (2005) and follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement [10], where applicable.

Results

Retrieved Studies

The database search retrieved 1478 citations; 607 duplicates were removed (Figure 1). After screening the abstracts and titles, 768 articles were excluded for not meeting the eligibility criteria. Full-text screening eliminated 36 articles (a list of

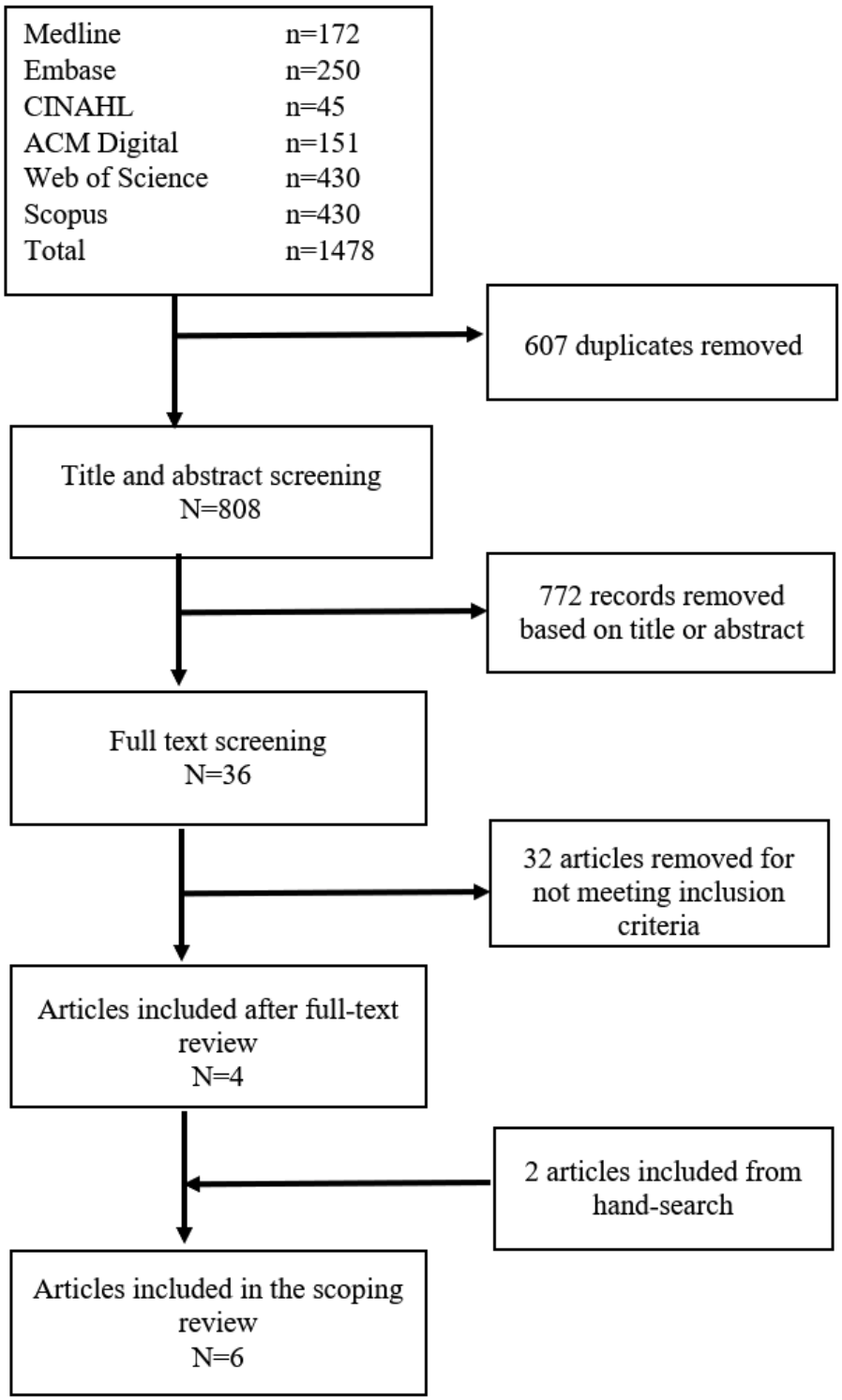
XSL•FO
RenderX

excluded articles is available in Multimedia Appendix 2). An additional 2 articles were found via hand-search, yielding 6 included studies in total. The kappa statistic measuring inter-rater agreement for full-text screening was 0.6 (moderate agreement) [11].

Description of Included Studies

The number of participants in studies ranged from 4 to 47 (Table 1). A total of 1 study was conducted in Canada [12], 2 in the United States [13,14], and 2 in Europe [15,16]. Most articles were published after 2010 except 1 [17]. The health domains covered in the studies included asthma [13], cardiovascular disease [16], kidney disease [12], Parkinson disease [17], diabetes [15], and mental health [14]. All included studies used quasi-experimental study designs to pilot test different context-aware systems. Demographic information on participants was often missing and inconsistently reported. Specifically, age data were reported by 3 studies [12,14,17], and 4 studies reported sex data [12,14,15,17].

Figure 1. Flow diagram of included studies.



XSL•FO
RenderX

Table 1. Characteristics of included studies and context-aware systems.

Study author, year, location	Health domain	Study type; duration	N total (mean age, % female)	Health activities	Patient-facing technologies	Functions
Bächlin et al, 2009, Israel [17]	Parkinson disease	Quasi-experimental; not reported (NR)	10 patients (66.4, 30)	Self-management of gait deficits in Parkinson patients	Acceleration sensors	Detect movement and freezing of gait
					Earphones and wearable computer	Produce sound when freezing of gait occurs (continuous external rhythmic auditory cues improve gait performance)>
Ong et al, 2016, Canada [12]	Chronic kidney disease	Quasi-experimental; 6 months	47 patients with chronic kidney disease (59, 47)	Self-management of chronic kidney disease (self-monitor blood pressure [BP] and symptoms, manage medications, track lab test results)	Wireless BP monitor	Measure BP
					Mobile app	Personalized real-time feedback on BP; reminders (eg, reconcile medication, and measure BP); self-monitor symptoms; access to lab test results and medication list
Lamprinos et al, 2016, Germany and Turkey [15]	Diabetes	Quasi-experimental; 6 weeks	In Germany: 21 patients (NR, 24); In Turkey: 39 patients (NR, 46)	Self-management of diabetes (self-monitor physiological measures; manage medications and lifestyle behaviors)	Mobile app and website	Self-monitor (eg, blood glucose, weight, BP, medication, physical activity, diet, and sleep); personalized feedback (decision making and action planning)
Zhang et al, 2016, Germany [16]	Cardiovascular disease	Quasi-experimental; NR	5 healthy young adults (NR)	Self-management of cardiovascular disease (self-monitor heart rate and identify abnormalities)	Wearable sensors	Track physical activity, heart rate, skin temperature, cardiac and pulmonary function, posture
					Environmental sensors	Detect room temperature
					Mobile app	Retrieve sensor data; trigger an alarm when an abnormal heartbeat is detected
Anantharam et al, 2015, United States [13]	Asthma	Quasi-experimental; 10 days	4 children (NR)	Self-management of asthma (self-monitor symptoms and identify triggers)	Indoor sensor	Monitor environmental and air quality observations (eg, pollen levels, carbon monoxide, temperature, and humidity)
					Exhaled air sensor	Monitor exhaled nitric oxide (indicator of inflammation)
					Mobile app	Gather and display sensor data; record users' observations (eg, asthma-related symptoms) via questionnaires; personalized feedback
Burns et al, 2011, United States [14]	Major depressive disorder	Quasi-experimental; 8 weeks	8 patients (37.4, 88)	Self-management of depression (self-monitor symptoms and identify triggers)	Mobile phone sensors	Collect data on location, ambient light, phone usage
					Website	Provide behavioral therapy; display data collected from the mobile phone
					Mobile app	Collect self-reported data on social context, activity, location, and internal states (ie, mood) via ecological momentary assessment; integrate self-reports with sensor data; personalized feedback; predict patient states based on self-reports and sensor data

Context Elements and Technologies of Current Interventions

The contextual elements of each included study are summarized in Table 2. A total of 2 studies focused on indoor settings [13,17], and 4 studies involved both indoor and outdoor settings [12,14-16]. Context was captured using sensors [12-14,16,17] and self-reports [12-15]. A total of 1 particular study used ecological momentary assessment to capture self-reported data on social context, activity, and internal states (ie, mood) [14]. Sensors captured data on location, ambient light [14], air quality [13], room temperature [16], and physiological measures (blood pressure [BP], heart rate, skin temperature, cardiac function, exhaled nitric oxide) [12,13,16]. Acceleration sensors were used to track movement and physical activity [16,17].

XSL•FO
RenderX

Table 2. Context elements present in included studies.

Study, year	Settings	Environmental features	User features	Utilization of context
Bächlin, 2009 [14]	Indoor	None	Movement tracking	Real-time movement tracking system triggering cueing sound upon detection of freezing of gait.
Ong, 2016 [12]	Indoor and outdoor	None	Blood pressure (BP)	Provide real-time personalized feedback on BP (eg, uncontrolled BP triggered reminder messages recommending an increase in frequency of self-monitoring).
Lamprinos, 2016 [15]	Indoor and outdoor	None	Physical activity tracking (step counts), sleep tracking, blood glucose, BP, weight, mood, nutrition	Creates a personalized action plan based on patient-recorded data and generates self-management recommendations.
Zhang, 2016 [16]	Indoor and outdoor	Room temperature	Physical activity tracking (standing, walking, running, jumping, walking upstairs or downstairs), heart rate, skin temperature, cardiac function, pulmonary function, posture	Trigger an alarm when an abnormal heartbeat is detected.
Anantharam, 2015 [13]	Indoor	Carbon monoxide, temperature, humidity, pollen levels	Exhaled nitric oxide, asthma-related symptoms (eg, coughing and chest tightness)	Provide personalized actionable recommendations based on sensor data and patient-reported information (eg, identify and alert patients regarding triggers).
Burns, 2011 [13]	Indoor and outdoor	Location sensing, ambient light	Physical activity tracking, social context (eg, interactions with other people), and internal states (mood, intensity of discrete emotions, fatigue, sense of accomplishment, concentration and engagement, and perceived control over current activities); manually self-reported via ecological momentary assessment	Predict patient states based on self-reported and sensor data (using machine learning), displaying them on the mobile app. Future iterations will involve the use of predicted states to provide real-time interventions.

Most studies used a combination of sensor technology and mobile apps, where the sensors collected context information, and the apps utilized those data to deliver personalized feedback [12-14,16]. Only 1 study did not involve the use of sensors, collecting contextual information solely through user self-reports [15]. In another intervention, a sensor was used without a mobile app, where context was harnessed with the help of a wearable computer, and earphones delivered auditory cues to improve gait performance in Parkinson patients [17].

Health Activities and Health-Related Measures

Self-monitoring was the most frequent health activity supported by context-aware systems in the included studies, where the collected data were then used to provide personalized feedback. Self-monitored data included physiological measures (eg, BP) [12,13,15,16], symptoms [12-14], and lifestyle behaviors [15]. Other health activities included tracking lab test results [12], managing medications [12,15], and practicing specific behaviors (eg, overcoming freezing of gait) [17].

Only 3 studies reported the impact of the intervention on health-related measures [12,14]. Specifically, Bächlin [17] found that the intervention had a sensitivity of 73.1% and a specificity of 81.6% in detecting freezing of gait events. Ong et al [12] found statistically significant reductions in home BP readings between baseline and after intervention (systolic BP: –3.4 mmHg; 95% CI –5.0 to –1.8; diastolic BP: –2.1 mmHg; 95% CI –2.9 to –1.2). Burns et al [14] found a significant decrease in self-reported depressive symptoms (P<.001; per-protocol Cohen d=3.43) and comorbid anxiety symptoms (P<.001, per-protocol Cohen d=2.58) [14]).

No studies mentioned a thorough evaluation of patient safety problems. A total of 4 studies highlighted technical issues such as system downtime [15], battery drainage problems [13,14], and wearable sensor issues in activity detection [16].

Discussion

Principal Findings

To the best of our knowledge, this is the first systematic scoping review to examine context-aware interventions to support patient work. The emerging nature of the field is reflected in the small number of included studies, their recent time of publication (all after 2010), and the predominance of quasi-experimental study designs. The majority of the papers involved pilot and usability testing with a small number of users; there were no randomized controlled trials (RCTs) to evaluate the efficacy of a context-aware system. In the included studies, context was captured using sensors or self-reports, sometimes involving both. Most studies used a combination of sensor technology and mobile apps to deliver personalized feedback. A total of 3 studies examined the impact of interventions on health-related measures, showing moderate-to-good sensitivity and specificity in detecting freezing of gait events in Parkinson patients [17], as well as significant improvements in BP [12] and reductions in depression symptoms and comorbid anxiety symptoms [14].

XSL•FO
RenderX

Comparison With Previous Literature

Other reviews have looked at the use of context awareness in health care [4,18]. Bricon-Souf (2007) found that there was a large gap between the requirements expressed by users and the context-aware prototypes developed. In addition, they reported that there was no consensus in the research community on how to model context and architectures to support its use. Similarly, Orwat et al revealed that most systems were described in their prototype stage and that implementation issues were rarely mentioned.

Our study, though only focused on patient work, also revealed comparable findings. Most studies described prototypes, and only 3 studies examined the impact of interventions on health-related measures, showing promising results in detecting freezing of gait events in Parkinson patients [17], as well as in improving BP [12] and depression and anxiety symptoms [14]. The use of context-awareness systems in patient work interventions has the potential to facilitate self-monitoring and improve the relevance and quality of the feedback provided, personalizing it to better fit participants' context [12-14,16]. This sort of "just-in-time" support [19] has the potential to facilitate patient work and improve the self-management of chronic conditions, by providing the advice patients need to make health management decisions at the right time, on a daily basis. Ameliorating self-management practices is a cornerstone of quality improvement efforts in chronic disease care and is associated with better health outcomes in several conditions such as type 2 diabetes [20].

The costs and risks of using context-aware systems for patient work were rarely reported in the included studies. A total of 4 studies highlighted technical issues such as system downtime [15], battery drainage problems [13,14], and wearable sensor issues in activity detection [16]. No studies mentioned a thorough evaluation of patient safety problems. Future studies should consistently report unintended effects and possible harms of the systems, such as privacy, technical issues, or any other unanticipated incidents [21].

Strengths and Limitations

This systematic scoping review has several strengths in terms of study design. First, an extensive search was performed across multiple databases to ensure that all relevant studies were captured. Second, the screening form was pretested and piloted before screening. Third, all full-text papers were screened by 2 independent reviewers. Finally, the kappa score of 0.6 for the full-text screening phase revealed an acceptable level of agreement.

The results of our study need to be interpreted in light of some limitations. Given that this is an emerging field in health informatics, there is a lack of longitudinal and experimental studies, which hampers the evaluation of the impact of these interventions. This is the reason why a systematic scoping review was conducted instead of a systematic review.

Another limitation was the exclusion of non-English papers. Even though this was conducted to ensure that all the authors could adequately understand and make an informed decision based on the abstracts, we might have missed important papers on patient work.

Implications for Research and Health Care

The use of context-aware systems to support patient work is a promising area of research, as these interventions have the potential to facilitate self-monitoring and provide personalized just-in-time feedback based on users' characteristics and environmental features, with the aim of improving disease management and clinical outcomes. Specifically, the increasing use of sensors to automatically collect context information could eliminate the need for self-reporting and manual data entry, streamlining the task of self-monitoring for chronic disease patients [22]. Furthermore, future applications of artificial intelligence have the potential to expand on the current capacity of these systems to provide personalized and relevant services to individuals [23], better supporting users with their health-related tasks, and decreasing the burden of patient work.

A common issue in context-aware systems research is the challenge of evaluating their real-world implementation [4,18]. Implementation fidelity is "the degree to which programs are implemented as intended by the program developers" [24]. It is known that implementation settings play a crucial role in the effectiveness of interventions, an issue that is at the core of implementation science's efforts to model the impact of context on outcomes [25]. To allow for implementation fidelity and replicability, studies of context-aware systems should describe the setting explicitly as well as provide sufficient details about the intervention and any potential adaptations for it to fit a different setting [26,27].

Conclusions

The use of context-aware systems to support patient work is an emerging area of research. RCTs are needed to evaluate the effectiveness of context-aware systems in improving patient work, self-management practices, and health outcomes in chronic disease patients. Future studies should consistently report the intervention and the settings in which the intervention is being implemented.

Acknowledgments

The authors would like to thank Mr Duy Hung Nguyen and Ms Keyu Wang for their assistance in the manuscript format. This research was supported by the National Health and Medical Research Council grant APP1134919 (Centre for Research Excellence in Digital Health) and Program Grant APP1054146.

XSL•FO
RenderX

Authors' Contributions

AYSL, SV, and EC contributed to the study design. HLT, LL, and AYSL contributed to the search strategy. HLT, LL, KY, and PM contributed to the screening. HLT, LL, and ABK were responsible for data extraction. HLT, LL, AYSL, KY, ABK, and EC wrote the first draft. All authors contributed to the revision and subsequent drafts.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 69KB - jmir_v21i6e10896_app1.pdf]

Multimedia Appendix 2

Excluded articles.

[PDF File (Adobe PDF File), 75KB - jmir_v21i6e10896_app2.pdf]

References

1. Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: connecting patients to context. J Am Med Inform Assoc 2015 Jan;22(1):2-10. [doi: 10.1136/amiajnl-2014-002826] [Medline: 25125685]
2. Hibbard JH. Moving toward a more patient-centered health care delivery system. Health Aff (Millwood) 2004;Suppl Variation:VAR133-VAR135. [doi: 10.1377/hlthaff.var.133] [Medline: 15471780]
3. Strauss A. Continual Permutations Of Action. New York: Aldine Transaction; 1993.
4. Bricon-Souf N, Newman CR. Context awareness in health care: a review. Int J Med Inform 2007 Jan;76(1):2-12. [doi: 10.1016/j.ijmedinf.2006.01.003] [Medline: 16488663]
5. Baldauf M, Dustdar S, Rosenberg F. A survey on context-aware systems. Int J Ad Hoc Ubiq Co 2007;2(4):263-277 [FREE Full text] [doi: 10.1504/IJAHUC.2007.014070]
6. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A context-sensing mobile phone app (Q sense) for smoking cessation: a mixed-methods study. JMIR Mhealth Uhealth 2016 Sep 16;4(3):e106 [FREE Full text] [doi: 10.2196/mhealth.5787] [Medline: 27637405]
7. 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 Dec 12;18(12):e321 [FREE Full text] [doi: 10.2196/jmir.6058] [Medline: 27956375]
8. Hébert ET, Stevens EM, Frank SG, Kendzor DE, Wetter DW, Zvolensky MJ, et al. An ecological momentary intervention for smoking cessation: the associations of just-in-time, tailored messages with lapse risk factors. Addict Behav 2018 Mar;78:30-35. [doi: 10.1016/j.addbeh.2017.10.026] [Medline: 29121530]
9. Dey AK, Abowd GD, Salber D. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. Hum Comput Interact 2009 Nov 12;16(2-4):97-166. [doi: 10.1207/S15327051HCI16234_02]
10. Liberati A, Altman D, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
11. Higgens JP, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. London: The Cochrane Collaboration; 2011.
12. Ong SW, Jassal SV, Miller JA, Porter EC, Cafazzo JA, Seto E, et al. Integrating a smartphone-based self-management system into usual care of advanced CKD. Clin J Am Soc Nephrol 2016 Jun 6;11(6):1054-1062 [FREE Full text] [doi: 10.2215/CJN.10681015] [Medline: 27173169]
13. Anantharam P, Banerjee T, Sheth A, Thirunarayan K, Marupudi S, Sridharan V. Knowledge-Driven Personalized Contextual mHealth Service for Asthma Management in Children. 2015 Presented at: IEEE International Conference on Mobile Services (MS); June 27-July 2, 2015; New York, NY, USA.
14. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
15. Lamprinos I, Demski H, Mantwill S, Kabak Y, Hildebrand C, Ploessnig M. Modular ICT-based patient empowerment framework for self-management of diabetes: design perspectives and validation results. Int J Med Inform 2016 Jul;91:31-43. [doi: 10.1016/j.ijmedinf.2016.04.006] [Medline: 27185507]

XSL•FO
RenderX

16. Zhang W, Thurow K, Stoll R. A context-aware mhealth system for online physiological monitoring in remote healthcare. Int J Comput Commun Contr 2016;11(1):142-156.
17. Bächlin M, Plotnik M, Roggen D, Giladi N, Hausdorff JM, Tröster G. A wearable system to assist walking of Parkinson s disease patients. Methods Inf Med 2010;49(1):88-95. [doi: 10.3414/ME09-02-0003] [Medline: 20011807]
18. Orwat C, Graefe A, Faulwasser T. Towards pervasive computing in health care - a literature review. BMC Med Inform Decis Mak 2008;8:26 [FREE Full text] [doi: 10.1186/1472-6947-8-26] [Medline: 18565221]
19. Nahum-Shani I, Smith SN, Spring BJ, Collins LM, Witkiewitz K, Tewari A, et al. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann Behav Med 2016 Sep 23:446-462. [doi: 10.1007/s12160-016-9830-8] [Medline: 27663578]
20. Tricco AC, Ivers NM, Grimshaw JM, Moher D, Turner L, Galipeau J, et al. Effectiveness of quality improvement strategies on the management of diabetes: a systematic review and meta-analysis. Lancet 2012 Jun 16;379(9833):2252-2261. [doi: 10.1016/S0140-6736(12)60480-2] [Medline: 22683130]
21. Coiera E, Ash J, Berg M. The unintended consequences of health information technology revisited. Yearb Med Inform 2016 Nov 10(1):163-169. [doi: 10.15265/IY-2016-014] [Medline: 27830246]
22. Kennedy CM, Powell J, Payne TH, Ainsworth J, Boyd A, Buchan I. Active assistance technology for health-related behavior change: an interdisciplinary review. J Med Internet Res 2012;14(3):e80 [FREE Full text] [doi: 10.2196/jmir.1893] [Medline: 22698679]
23. Isern D, Moreno A. A systematic literature review of agents applied in healthcare. J Med Syst 2016 Feb;40(2):43 [FREE Full text] [doi: 10.1007/s10916-015-0376-2] [Medline: 26590981]
24. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implement Sci 2007 Nov 30;2:40 [FREE Full text] [doi: 10.1186/1748-5908-2-40] [Medline: 18053122]
25. Helfrich C, Damschroder L, Hagedorn H, Daggett G, Sahay A, Ritchie M, et al. A critical synthesis of literature on the promoting action on research implementation in health services (PARIHS) framework. Implement Sci 2010 Oct 25;5:82 [FREE Full text] [doi: 10.1186/1748-5908-5-82] [Medline: 20973988]
26. Cotterill S, Knowles S, Martindale A, Elvey R, Howard S, Coupe N, et al. Getting messier with TIDieR: embracing context and complexity in intervention reporting. BMC Med Res Methodol 2018 Jan 18;18(1):12 [FREE Full text] [doi: 10.1186/s12874-017-0461-y] [Medline: 29347910]
27. Coiera E, Ammenwerth E, Georgiou A, Magrabi F. Does health informatics have a replication crisis? J Am Med Inform Assoc 2018 Apr 13:963-968. [doi: 10.1093/jamia/ocy028] [Medline: 29669066]

Abbreviations

BP: blood pressure
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 27.04.18; peer-reviewed by N Archer, C Granja; comments to author 25.08.18; revised version received 09.04.19; accepted 26.04.19; published 17.06.19.

Please cite as:
Yin K, Laranjo L, Tong HL, Lau AYS, Kocaballi AB, Martin P, Vagholkar S, Coiera E
Context-Aware Systems for Chronic Disease Patients: Scoping Review
J Med Internet Res 2019;21(6):e10896
URL: http://www.jmir.org/2019/6/e10896/
doi:10.2196/10896
PMID:31210138

©Kathleen Yin, Liliana Laranjo, Huong Ly Tong, Annie YS Lau, A Baki Kocaballi, Paige Martin, Sanjyot Vagholkar, Enrico Coiera. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 17.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Social Media and Social Functioning in Psychosis: A Systematic Review

Jone Bjornestad[1], MSc, PhD; Wenche Ten Velden Hegelstad[2], MSc, PhD; Henrik Berg[3], MSc, PhD; Larry Davidson[4,5], MSc, PhD; Inge Joa[2,6], MSc, PhD; Jan Olav Johannessen[2,6], MD, PhD; Ingrid Melle[7], MD, PhD; Helen J Stain[8], MSc, PhD; Ståle Pallesen[9], MSc, PhD

[1]Department of Social Studies, Faculty of Social Sciences, University of Stavanger, Stavanger, Norway

[2]Network for Clinical Research in Psychosis, Stavanger University Hospital, Stavanger, Norway

[3]Norsk Lærer Akademi, University College, Bergen, Norway

[4]School of Medicine, Yale University, New Haven, CT, United States

[5]Institution for Social and Policy Studies, Yale University, New Haven, CT, United States

[6]Network for Medical Sciences, Faculty of Health, University of Stavanger, Stavanger, Norway

[7]Norwegian Centre for Mental Disorders Research, Faculty of Medicine, University of Oslo, Oslo, Norway

[8]School of Social and Health Sciences, Leeds Trinity University, Leeds, United Kingdom

[9]Department of Psychosocial Science, University of Bergen, Bergen, Norway

Corresponding Author:
Jone Bjornestad, MSc, PhD
Department of Social Studies
Faculty of Social Sciences
University of Stavanger
PO Box 8600 Forus
Stavanger, 4036
Norway
Phone: 47 97141599
Email: jone.r.bjornestad@uis.no

Abstract

Background: Individuals with psychosis are heavy consumers of social media. It is unknown to what degree measures of social functioning include measures of online social activity.

Objective: To examine the inclusion of social media activity in measures of social functioning in psychosis and ultrahigh risk (UHR) for psychosis.

Methods: Two independent authors conducted a search using the following electronic databases: Epistemonikos, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, MEDLINE, Embase, and PsycINFO. The included articles were required to meet all of the following criteria: (1) an empirical study published in the English language in a peer-reviewed journal; (2) the study included a measure of objective or subjective offline (ie, non-Web-mediated contact) and/or online social functioning (ie, Web-mediated contact); (3) the social functioning measure had to be used in samples meeting criteria (ie, Diagnostic and Statistical Manual of Mental Disorders or International Classification of Diseases) for a psychotic disorder or UHR for psychosis; and (4) the study was published between January 2004 and February 2019. Facebook was launched as the first large-scale social media platform in 2004 and, therefore, it is highly improbable that studies conducted prior to 2004 would have included measures of social media activity.

Results: The electronic search resulted in 11,844 distinct articles. Full-text evaluation was conducted on 719 articles, of which 597 articles met inclusion criteria. A total of 58 social functioning measures were identified. With some exceptions, reports on reliability and validity were scarce, and only one measure integrated social media social activity.

Conclusions: The ecological validity of social functioning measures is challenged by the lack of assessment of social media activity, as it fails to reflect an important aspect of the current social reality of persons with psychosis. Measures should be revised to include social media activity and thus avoid the clinical consequences of inadequate assessment of social functioning.

XSL•FO
RenderX

Trial Registration: International Prospective Register of Systematic Reviews (PROSPERO) CRD42017058514; http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017058514

(J Med Internet Res 2019;21(6):e13957) doi:10.2196/13957

KEYWORDS

psychosis; schizophrenia; social media; social functioning; measures; assessment; systematic review

Introduction

Social functioning impairment is a core dimension of psychotic disorders [1]. Thus, measures of social functioning are crucial for clinical assessment, prognosis, and outcome. Research indicates high engagement with social media platforms and associated social interaction in subjects with psychosis and those at ultrahigh risk (UHR) for psychosis states, including friendship formation and overcoming barriers associated with having severe psychiatric symptoms [2-5]. Social media activity should therefore be included as part of social functioning measures.

In an Australian national survey, more than one-third of adults with psychosis rated social functioning issues as their greatest challenge for the future [6]. Long-term deficits in social functioning have been linked to negative symptoms, such as social withdrawal, apathy, and avolition [7,8], as well as impaired social cognitive capacities, including capacity for mentalization and theory of mind [5,9]. Similar findings have been found for UHR populations; when compared to healthy controls, they show both higher levels of baseline social decline and lower levels of quality of life [10-14]. Conversely, good social functioning has been identified as a robust predictor of recovery [15-18].

Empirical research on social functioning largely originates from standardized questionnaires based on two dimensions [19]. The objective dimension encompasses the ability to meet social roles, such as employability and being a spouse, a family member, or a friend, combined with socioeconomic measures, such as finances and housing [20]. These measures are easily quantifiable and can thus be replicated [20]. The subjective dimension comprises self-reported measures of social roles and measures of satisfaction with family life, recreational activities, and life as a whole [20]. Ratings on both objective and subjective measures are found to correlate with prognosis, course development, and outcome [21].

Since the advent of Facebook in 2004, social media is exponentially more often involved in establishing and maintaining social networks [22-24]. Globally, there are approximately 2.6 billion registered social media profiles and the number is expected to grow by an additional 400 million over the next three years [25]. In 2015, in the United States, more than 75% of people used social media compared to 7% a decade ago, and 92% of adolescents went online daily [26]. Nonetheless, the conceptualization of social media participation as a dimension of social functioning is underdeveloped. At face value, when compared with offline contact, social media platforms represent radically evolving platform structures and a more asynchronistic kind of communication. These are technology-mediated tools that enable individuals to share, exchange, and create ideas, images, and information through online communities and networks [27-29].

Despite having fewer or less-frequent social contacts outside social media, individuals with psychosis or those at UHR for psychosis are heavy consumers of social media when compared to peers of the same age [30-34]. The Internet has become an influential source of mental health information for people with psychosis [28] and, thus, social media and digital devices have been utilized to support mental health care [32-35] and destigmatization campaigns [36]. Particularly for the youngest age group with psychosis and those at UHR for psychosis, there has already been social media-based interventions developed that are targeted on psychological, functional, and social recovery [37,38].

Science and technology studies aim at offering a comprehensive understanding of the interaction between science, technology, and society [39]. According to this framework, technologies may fundamentally alter societal dynamics, influencing communication. Moreover, post-normal science (PNS) is a perspective emphasizing the value of direct stakeholder involvement in practices where facts are uncertain and stakes are high [40], as they arguably are in psychosis. If measures of social functioning in psychosis do not embody the fundamental changes caused by technological innovations and do not consult target groups directly, they run the risk of low ecological validity.

The main objective of this study was to examine whether measures of social functioning in psychosis and UHR for psychosis include the assessment of social behavior on social media. It also compared the validity and reliability of reported measures of social functioning.

Methods

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [41] to ensure comprehensive and transparent reporting of methods and results. The protocol was registered at the International Prospective Register of Systematic Reviews (PROSPERO) in March 2017 (registration number: CRD42017058514).

Search Strategy

Two independent authors (JB and WTVH) conducted a search using the following electronic databases: Epistemonikos, Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects (DARE), MEDLINE, Embase, and PsycINFO. The search terms used were as follows: (“psychosis” or “psychoses” or “psychotic*” or “schizo*”) AND (“social*” or “psychosocial*” or “communit*” or “peer*” or

XSL•FO
RenderX

"famil*" or "friend*"). Specific search terms were added to capture social media activity (eg, the Medical Subject Headings [MeSH] term "social media"; see Multimedia Appendix 1 for model search). The search queries were reviewed by an information scientist and were limited to title, abstract, keywords, and subject headings. In addition, a manual literature search was performed using reference lists of reviews and meta-analyses. In cases of doubt, the full-text paper was read to determine eligibility. Papers published between January 2004 and February 2019 were included. The last search was conducted on February 15, 2019.

Eligibility Criteria

The included articles were required to meet all of the following criteria:

1. Empirical study published in the English language in a peer-reviewed journal.
2. The study included a measure of objective or subjective offline (ie, non-Web-mediated contact) and/or online social functioning (ie, Web-mediated contact).
3. The social functioning measure had to be used in samples meeting criteria (ie, Diagnostic and Statistical Manual of Mental Disorders [DSM] or International Classification of Diseases [ICD]) for a psychotic disorder or UHR for psychosis.
4. The study was published between January 2004 and February 2019. Facebook was launched as the first large-scale social media platform in 2004 and, therefore, it is highly improbable that studies conducted prior to 2004 would have included measures of social media activity.

Exclusion Criteria

Articles were excluded if the only functioning assessed by the measure was one of the following:

1. Premorbid functioning measures.
2. Global functioning measure.
3. Performance-based skills assessment.
4. Studies exclusively dealing with social relationships, including social functioning, between study participants and professionals.

Data Collection

All potential studies were exported into a reference citation manager, EndNote (Clarivate Analytics), before removing duplicates. Two independent reviewers (JB and WTVH) separately performed the screening of titles and abstracts, full-text analysis, and selection of social functioning measures. Disagreements were resolved through discussion until consensus was reached. A third reviewer (SP) was available to resolve disagreements. Finally, the list of included and excluded measures was sent to an independent auditor (HJS) for critical evaluation. The kappa coefficient was used to assess the level of agreement of the two independent reviewers for the selection of included and excluded measures.

Analytic Methods and Data Extraction Procedure

A narrative descriptive synthesis was performed for the included articles. The data extraction procedure was performed in two steps. First, subjective and objective measures of social functioning across different social domains (ie, work, community functioning, socioeconomic status, etc) for both offline and online engagements were identified. Second, the content, quality, and psychometric properties, with a particular focus on whether measures assessed social media activities and interactions, were examined and assessed, including validity and reliability statistics of the measures. Since the selection of screened and included articles was extensive, validation literature was sourced directly from the reviewed articles. In addition, a manual search was performed for each individual measure to identify further validation literature.

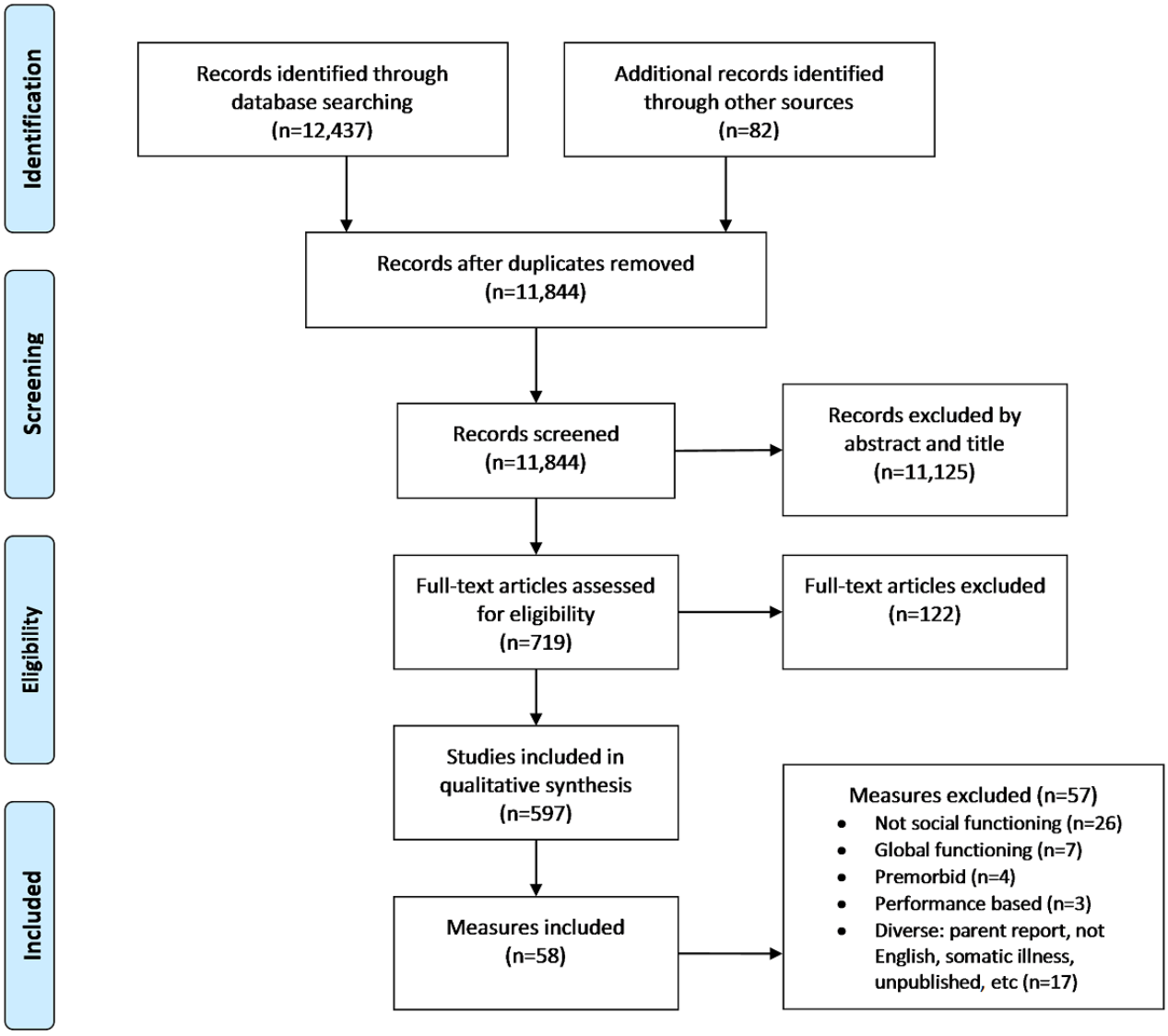
Results

Search Results

The electronic search returned 12,437 articles. After duplicates were removed, there were 11,844 articles, of which 671 were reviews or meta-analyses: 178 articles from the Cochrane Database of Systematic Reviews and 493 from Epistemonikos. A hand search of reference lists of reviews and meta-analyses returned a further 82 articles. Full-text evaluation was conducted for 719 articles, of which 597 met the inclusion criteria and were included for the final synthesis. From the 597 articles, 58 measures were identified: 41 (71%) social functioning and 17 (29%) quality-of-life measures that included assessment of social functioning. Interrater reliability (ie, agreement between independent reviewers) for inclusion of measures was high (k=.83). Details of the search results are summarized in Figure 1.

XSL•FO
RenderX

Figure 1. Flow diagram of the reviewing process according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.



Frequently Used Social Functioning Measures

Details of the included 58 measures of social functioning are summarized in Tables 1 and 2. The three most frequently used measures were the Social Functioning Scale (78 references), the Quality-of-Life Scale (67 references), and the World Health Organization Quality of Life Brief Version (WHOQOL-BREF) (57 references). Several measures (eg, the Social Adjustment Scale II and the Global Functioning-Social Scale) had been used to address social functioning in UHR populations. Although developed for young people, none of these measures were exclusively used in UHR populations.

XSL•FO
RenderX

Table 1. Included social functioning measures (N=41).

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Assessment of Communication and Interaction Skills	Forsyth et al, 1999, 1	Observer-rated 20 items, 4-point Likert response format Three domains: physicality, information exchange, and relations Gathers data on skill as it is exhibited during performance in an occupational form and/or within a social group context	Psychosis and general psychiatry	Good internal and construct validity [43,44] Good discriminant validity [43,44]	Good interrater reliability [43,44]	No
Behavior and Symptom Identification Scale	Eisen et al, 1999, 1	Self-report 32 items, 4-point Likert response format Five domains: relationship (self and others), daily living and role performance, depression and anxiety, impulsive and addictive behavior, and psychosis	Psychosis, inpatients and outpatients	Questionable internal and construct validity; poor for psychosis [45] Good discriminant validity; unacceptable for psychosis [45] Good sensitivity to change [45]	Good interrater reliability [45]	No
Client's Assessment of Strengths, Interests and Goals	Wallace et al, 2001, 2	Self-report and informant-report 102 items, 5-point Likert response format Six domains: current social and independent living skills, medication compliance and side effects, quality of life, quality of treatment, symptoms, and performance of unacceptable community behaviors	Psychosis and general psychiatry	Questionable concurrent validity Good discriminant validity [46]	Good test-retest reliability Questionable interrater reliability Good internal consistency [46]	No
Community Adjustment Form	Stein & Test, 1980, 7	Observer-rated semistructured interview 140 items, 10-item scale 17 domains (eg, living situation, work and social functioning, family involvement, and medication use) Also includes an observer-rated, 10-item scale of prosocial behaviors	Psychosis	No data available	Excellent interrater reliability [47]	No
Disability Assessment Schedule—II	WHO[c], 2010, 31	Observer-rated 12, 36, or 97 items; scoring based on all available information (eg, patient´s written records or data from informants) Six domains: understanding and communicating, getting around, self-care, getting along with others, household and work activities, and participation in society Several versions: DAS[d], DAS-II-sv[e], SDSS[f], and WHO-DAS[g]	General psychiatry	Moderate concurrent validity [48]	Good internal consistency [48] Excellent test-retest reliability [48]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
First Episode Social Functioning Scale	Lecomte et al, 2014, 5	Self-report 34 items, 4-point Likert response format Eight domains: independent living skills, interacting with people in different contexts, social activities, intimacy, friendships, family relations, work, and school Perceived ability and actual behavior rated for each item	Psychosis	Good convergent validity [42] Good discriminant validity [42]	Good sensitivity to change [42]	Yes
Functional Assessment Short Test	Rosa et al, 2007, 1	Self-report 24 items, 4-point Likert response format Six domains: autonomy, occupational functioning, cognitive functioning, financial issues, interpersonal relationships, and leisure time	Bipolar disorder	Good concurrent validity Questionable discriminant validity [49]	Good test-retest reliability [49] High internal consistency [49]	No
Functional Remission of General Schizophrenia	Llorca et al, 2009, 5	Observer-rated 19 items, 5-point Likert response format Six domains: daily life, activities, relationships, quality of adaptation, health, and treatment	Schizophrenia	Good concurrent validity [50]	Good internal consistency [50]	No
Global Assessment of Relational Functioning	Dausch BM et al, 1996, 1	Observer-rated 20 domains of relationship functioning, 100-point scale (81-100 indicates satisfying functioning)	Bipolar disorder	Good concurrent validity Good discriminant validity [51]	Good-to-high interrater reliability [51] Good internal consistency [51]	No
Global Functioning—Social	Cornblatt B et al, 2007, 22 (5 UHR[h])	Observer-rated Seven probe questions, 10-point Likert response format Assesses levels of social contact and friendships outside of the family	Psychosis	Good construct validity [52]	Good interrater reliability [52]	No
Groningen Social Disability Schedule	Wiersma, 1988, 8	Observer-rated semistructured interview 5-point Likert response format Eight domains: self-care, relationship with the family, relationship with a partner, friendship, parental role, citizenship, leisure activities, and work and occupation	General psychiatry	Poor-to-excellent sensitivity to change [53] Good discriminant validity [54]	Good interrater reliability Questionable test-retest reliability [53]	No
Health of the Nation Outcome Scale	Wing et al, 1998, 8	Observer-rated 12 items, 4-point Likert response format Aggression, self-harm, drug and alcohol use, cognitive problems, physical illness and disability, hallucinations and delusions, depression, other symptoms, social relationships, activities of daily living, residential environment, and day-time activity	General psychiatry	Good concurrent validity [55] Good discriminant validity [55]	Poor-to-good interrater reliability [55] Poor-to-acceptable test-retest reliability [55]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Index of Social Competence	McConkey R & Walsh J, 1982, 1	Observer-rated 15 items of ability Six domains: community skills, self-care skills, communication skills, time, money, and additional handicaps	General psychiatry	No data available	Good interrater reliability [56]	No
Inventory of Interpersonal Problems	Horowitz et al, 1988, 1	Self-report 32 items, 4-point Likert response format Nine domains: domineering or controlling, vindictive or self-centered, sociable, intimate, submissive, responsible, nonassertive, self-sacrificing, and intrusive	Community: nonclinical	No data available	Acceptable-to-good test-retest reliability [57] Good internal consistency [57]	No
Interview Schedule for Social Interaction	Henderson S et al, 1980, 2	Observer-rated 53 items, individual summary scores for each domain Four domains: availability of close relationships, adequacy of above, availability of friendships, and adequacy of above	General population	Good face validity [58]	Acceptable test-retest reliability [58] Acceptable internal consistency [58]	No
Life Skills Profile	Parker et al, 1991, 14	Observer-rated 39 items, 4-point Likert response format Five domains: ability for self-care, turbulent behavior (reverse scored), sociability, communication, and responsibility	Psychosis, inpatients and outpatients	Questionable construct validity [59]	Good interrater reliability [59] Good test-retest reliability [59]	No
Multnomah Community Ability Scale	Barker et al, 1994, 12	Observer-rated 17 items, 5-point Likert response format Four domains: assessment of interference with functioning, adjustment to living, social competence, and behavioral problems	General psychiatry, chronic patients	No data available	Good interrater reliability [60] Good test-retest reliability [60]	No
Personal and Social Performance	Morosini et al, 2000, 55	Observer-rated Single-item, 100-point response format (score determined by domain score range) Four domains (6-point response format per domain): socially useful activities, personal and social relationships, self-care, disturbing and aggressive behaviors	Psychosis	Good construct validity [61]	Good interrater reliability [61] Excellent test-retest reliability [61]	No
Provision of Social Relations Scale	Turner et al, 1983, 2	Self-report 15 items that measure social relationships with family and friends, 5-point Likert response format	Schizophrenia, bipolar disorder, and healthy controls	No data available	Good test-retest reliability [62]	No
Psychosocial Functioning Scale	Valencia et al, 1989, 1	Self-report 35 items, 5-point Likert response format Five domains: occupational, social, money management, marital, and familial	General psychiatry	Good construct validity	Good internal consistency [63]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Recovery Assessment Scale	Corrigan et al, 1999, 2	Self-report 41 items, 5-point agreement scale Four domains: doing things I value, looking forward, mastering my illness, and connecting and belonging	General psychiatry, chronic patients	Good concurrent validity [64]	Good test-retest reliability [64]	No
Rehabilitation Evaluation Hall and Baker	Baker R & Hall JN, 1988, 1	Observer-rated 22 items, 9-point Likert response format Two domains: deviant behavior and general behavior General Behavior subscale (only subscale relevant for the cited study): 15 items Five domains: social activity, speech disturbance, speech skills, self-care skills, and community skills	General psychiatry	Good criterion validity [65] Good discriminant validity [65]	Good interrater reliability [65]	No
Role Functioning Scale	Goodman et al, 1993, 7	Observer-rated Four items, 7-point Likert response format Four domains: work, independent living and self-care, immediate social network relationships, and extended social network relationships	Psychosis and depression	Good discriminant validity [66]	Poor-to-good interrater reliability [66] Good test-retest reliability [66] Good internal consistency [66]	No
Schizophrenia Social Functioning Index	Padmavathi R, 1995, 2	Observer-rated 17 items, 5-point Likert response format Four domains: self-concern, occupational role, role in family, and other social roles Each section has several subsections covering different areas of social functioning	Schizophrenia and relatives	Good concurrent validity [67]	Good interrater reliability [67] Good test-retest reliability [67] Acceptable internal consistency [67]	No
Self Evaluation and Social Support—Schizophrenia version	Humphreys et al, 2001, 1	Observer-rated Five domains: social and recreational, occupational, relationships, parenting, and homemaking Questions in these sections involve both perceived competence and commitment in each possible role, and responses are used in an overall rating; also, sections covering self-evaluation of personal attributes and self-acceptance	Psychosis	No data available	Acceptable-to-good interrater reliability [68] Poor test-retest reliability [68]	No
Short Form Health Survey—36	Ware & Donald, 1992, 12	Observer-rated or self-report 36 items, 100-point response format Eight domains: physical functioning, role limitations due to physical health problems, bodily pain, general health perceptions, vitality, social functioning, role limitation due to emotional problems, and general mental health	Psychosis	Moderate discriminant validity [69]	Good test-retest reliability [69]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Social Adaptation Self-Evaluation Scale	Bosc M et al, 1997, 4	Self-report 21 items, 4-point Likert response format Four domains: social, familial, occupational, and environmental functioning	Major depression	Acceptable sensitivity to change [70]	Poor test-retest reliability [70] Good internal consistency [70,71]	No
Social Adaptive Functioning Evaluation	Harvey et al, 1997, 1	Observer-rated 17 items, 5-point Likert response format Three domains: impulse control, instrumental and self-care, and social functions	Schizophrenia	Good convergent validity [72] Good discriminant validity [72]	Good interrater reliability [72] Good test-retest reliability [72]	No
Social Adjustment Inventory for Children and Adolescents	Gammon et al, 1987, 1 (UHR)	Observer-rated interview Multiple domains: functioning in school, spare time activities, and interactions with peers and family	Children and adolescents of parents with and without major depression	Questionable-to-poor convergent validity [73]	Poor interrater reliability [71,73] Good-to-poor internal consistency [73]	No
Social Adjustment Scale II	Paykel et al, 1971, 27 (1 UHR)	Observer-rated semistructured interview 52 items, 5-point Likert response format Eight domains: work, domestic relationship, parental role, relationship with external family, social and leisure activities, sexual activity, romantic involvement, and personal well-being Self-report version [74]	Depression	Poor-to-acceptable convergent validity [74]	No data available	No
Social Behaviour Scale	Wykes & Sturt, 1986, 16	Observer-rated 21 items, 5-point Likert response format Six domains: occupation, behavioral problems, personal self-care, leisure activities, performance and expectations, and communication skills	General psychiatry	Good discriminant validity [75] Good concurrent validity [75]	Good interrater reliability [76] Good interinformant reliability [76] Good test-retest reliability [76]	No
Social Functioning Questionnaire	Tyrer et al, 2004, 1	Self-report Eight-item assessment of perceived social functioning, score 0-24 Developed from the Social Functioning Scale	Psychiatric outpatient (nonpsychotic)	No data available	Good interrater reliability [77] Good test-retest reliability [77] Good internal consistency [77]	No
Social Functioning Scale	Birchwood et al, 1990, 78	Self-report or observer-report 79 items, 4-point Likert response format Seven domains: social engagement, interpersonal behavior, prosocial activities, recreation, independence-competence, independence-performance, and employment and occupation	Psychosis	Good construct validity [78] Good discriminant validity [78] Good convergent validity [78]	Good interrater reliability [78]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Social Integration Survey	Kawata AK & Revicki DA, 2008, 1	Self-report or informant-report 62 items, 4-6-point Likert response format Five domains: social perception, work interactions, social skills, social cognition, and daily living skills or self-care	Schizophrenia	Good discriminant validity [79]	Good internal consistency [79] Poor interrater reliability [79]	No
Social and Occupational Functioning Assessment Scale	American Psychiatric Association, 1994, 35	Observer- or self-report Single-item, 100-point response form	General psychiatry	Acceptable sensitivity to change [80]	No data available	No
Social Occupational Functioning Scale	Saraswat N, 2006, 2	Observer-rated 15 items, 5-point Likert response format Three domains: adaptive living skills, social appropriateness, and interpersonal skills	Schizophrenia	Acceptable concurrent validity [81] Acceptable criterion validity (positive and negative symptom total score) [81] Acceptable discriminant validity [81]	High internal consistency [81] Good test-retest [81]	No
Social Role Functioning Test	McPheeters H, 1984, 1	Observer-rated 28 items, 7-point Likert response format Three domains: work productivity, independent living, and social network relationships (immediate and extended)	General psychiatry	No data available	Good internal consistency [82]	No
Specific Levels of Functioning	Schneider & Struening, 1983, 19	Observer-rated 43 items, 5-point Likert response format Six dimensions: physical functioning, personal care skills, interpersonal relationships, social acceptability, activities, and work skills	Psychosis and general psychiatry	No data available	Excellent internal consistency [83] Excellent interrater reliability [83]	No
Strauss Carpenter Level of Functioning	Strauss & Carpenter, 1977, 10	Observer-rated Four items, 5-point Likert response format Four domains: symptomatology, work, social contacts, and function	Not stated	No data available	No data available	No
Strauss Carpenter Outcome Scale	Strauss & Carpenter, 1972, 7	Observer-rated Four items, 49-point Likert response format Four domains: social activities, work, independent living, and hospitalization	Schizophrenia	No data available	No data available	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Time Budget Measure	Jolley et al, 2006, 1	Observer-rated Diary-based measure (28 time blocks for the week), score range 0-112	Psychosis	Acceptable convergent validity [84]	Good interrater reliability [85] Good test-retest reliability [85]	No

[a]Number of citations the scale has gotten throughout the years, which indicates the scale popularity and impact.

[b]Reliability: 1 (perfect reliability), ≥.9 (excellent reliability), ≥.8<.9 (good reliability), ≥.7<.8 (acceptable reliability), ≥.6<.7 (questionable reliability), ≥.5<.6 (poor reliability), <.5 (unacceptable reliability), 0 (no reliability).

[c]WHO: World Health Organization.

[d]DAS: Disability Assessment Schedule.

[e]DAS-II-sv: Disability Assessment Schedule—II: Schizophrenia Version

[f]SDSS: Social Disability Screening Schedule.

[g]WHO-DAS: World Health Organization Disability Assessment Schedule

[h]UHR: ultrahigh risk.

XSL•FO
RenderX

Table 2. Included quality-of-life measures (N=17).

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Assessment of Quality of Life	Hawthorne et al, 1999, 2	Self-report 47 items, 4-point Likert response format Five domains: illness, independent living, social relationships, physical senses, and psychological well-being	Clinical and community sample	Good discriminant validity [86]	Good internal consistency [86]	No
Health Related Quality of Life	Nelson et al, 1987, 3	Self-report 5-point Likert response format Nine global questions, each illustrated with drawings to measure the following domains: physical fitness, pain, feelings and emotions, daily activities, social activities, change in health, overall health, social support, and overall quality of life.	Mainly chronic illnesses	Good face validity [87]	Good test-retest reliability [88] Good interrater reliability [88]	No
Lancashire Quality of Life Profile	Oliver, 1991, 3	Observer-rated 100 items, 7-point Likert response format 10 domains: living situation, leisure and social participation, health, finances, family relations, safety, positive esteem, negative esteem, framework, and fulfilment	General psychiatry	Moderate-to-good concurrent validity [89]	Moderate-to-good internal consistency [89] Good test-retest reliability [89]	No
Manchester Short Assessment of Quality of Life	Priebe et al, 1989, 23	Observer-rated clinical interview 25 items, 7-point Likert response format 12 domains with three subscales: these include stable personal patient details, personal details that may change over time (eg, education), and questions that must be asked at each assessment, including both objective and subjective items concerning quality of life and social life	Psychosis and students	Good concurrent validity [90]	Good internal consistency [91]	No
Modular System for Quality of Life	Pukrop R et al, 2000, 3	Self-report 47 items, 7-point Likert response format Seven domains: physical health; vitality; and psychosocial, affective, material, spare time, and general quality of life	Schizophrenia and general population	Poor discriminant validity [92]	Good internal consistency [92] Good test-retest reliability [92]	No
Sickness Impact Profile	Pollard et al, 1976, 1	Self-report 48 items (short version), individual category scores Four domains: sleep and rest, home management, contact with family and friends, and leisure activities	Psychiatric samples and somatic samples	High internal consistency [93] Good concurrent validity [94]	High test-retest reliability [93] High interrater reliability [93]	No
Quality of Life Enjoyment and Satisfaction Questionnaire—Short Form	Endicott et al, 1993, 8	Observer-rated 59 items, 5-point Likert response format Five domains: general activities, physical activities, emotional functioning, recreational activities, and social relationships	Psychosis and depression	Poor-to-moderate discriminant validity [95]	Good internal consistency [96] Moderate test-retest reliability [96,97]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Quality of Life Index	Ferrans CE & Powers MJ, 1985, 1	Self-report 32 items, 6-point Likert response format Four domains: health and functioning, social and economic, psychological and spiritual, and family	Schizophrenia and healthy controls	Good convergent validity [98]	Excellent internal consistency [98]	No
Quality of Life Interview	Lehman, 1983, 29	Observer-rated semistructured interview 143 items (brief versions: 33 or 78 items) Eight domains: accommodation, family, social relations, leisure, safety, finances, physical health, and mental health	Psychosis and general psychiatry	Good construct validity [99]	Good-to-acceptable internal consistency [99] Poor-to-good test-retest reliability [99]	No
Quality of Life Inventory—74	Frisch MB, 1992, 2	Self-report 17 items, 4-point Likert response format	General psychiatry, undergraduates, and forensic	Good convergent validity [100] Good construct validity [100]	Good test-retest reliability [100] Good internal consistency [100]	No
Quality of Life Scale	Heinrichs et al, 1984, 67	Observer-rated semistructured interview 21 items, 7-point Likert response format Four domains: intrapsychic foundations, interpersonal relations, instrumental role, and common objects and activities	Psychosis and general psychiatry	Poor-to-good convergent validity [101]	Excellent inter-rater reliability [101]	No
Quality of Well-Being	Kaplan et al, 1978, 2	Observer-rated Preference weight for each domain Three domains: mobility, physical activity, and social activity	Chronic somatic illness	Good discriminant validity [102,103] Good convergent validity [102,103]	No data available	No
Satisfaction With Life Scale	Test et al, 2005, 6	Self-report 18 items, 5-point Likert response format Four domains: living situation, work, social relationships, and self and present life	Schizophrenia and general psychiatry	Good construct validity [104] Good concurrent validity [104]	Good internal consistency [104]	No
Schizophrenia Quality of Life—18	Boyer et al, 2010, 1	Self-report 41 items, index scores from 0 to 100 Eight domains: psychological well-being, self-esteem, family relationships, relationships with friends, resilience, physical well-being, autonomy, and sentimental life	Schizophrenia	Good-to-unacceptable concurrent validity; great scale variability [105]	Good-to-acceptable internal consistency [105] Good-to-acceptable test-retest reliability [105]	No
Schizophrenia Quality of Life Scale	Wilkinson et al, 2000, 10	Self-report 30 items, 5-point Likert response format Three domains: psychosocial, motivation and energy, and symptoms and side effects	Schizophrenia	Good construct validity [106]	Good internal consistency [106] Good test-retest [106]	No

XSL•FO
RenderX

Measure	Author, year, number of scale citations[a]	Description	Validation samples	Validity	Reliability[b]	Social media (Yes or No)
Wisconsin Quality of Life Index	Becker, 1993, 4	Self-report 113 items, 5-point Likert response format Nine domains: general life satisfaction, activities and occupations, psychological well-being, physical health, social relations and support, economics, activities of daily living, symptoms, and goal attainment	Schizophrenia and general psychiatry	Good convergent validity [107]	Good-to-acceptable internal consistency [107] Good test-retest reliability [107]	No
WHOQOL-BREF[c]	WHO Quality of Life Group, 1998, 57	Self-report 268 items, 5-point Likert response format Four domains: physical, psychological, social, and environmental	Psychosis and general psychiatry	Poor-to-moderate construct validity [108]	Good internal consistency [108]	No

[a]Number of citations the scale has gotten throughout the years, which indicates the scale popularity and impact.

[b]Reliability: 1 (perfect reliability), ≥.9 (excellent reliability), ≥.8<.9 (good reliability), ≥.7<.8 (acceptable reliability), ≥.6<.7 (questionable reliability), ≥.5<.6 (poor reliability), <.5 (unacceptable reliability), 0 (no reliability).

[c]WHOQOL-BREF: World Health Organization Quality of Life Brief Version.

Structure and Administration of Measures

A total of 35 out of 58 measures (60%) were primarily observer-rated, while 23 (40%) were primarily self-reported. The completion time ranged from 10 minutes (ie, Social Functioning Questionnaire) to 60 minutes (ie, Social Adjustment Scale). Most of the social functioning and quality-of-life measures used a Likert response format (40/58, 69%). Most measures assessed behaviors, not perceived ability, related to physical forms of social functioning, such as face-to-face or telephone contact with friends and family. There was great variability in how comprehensive measures reported on social functioning characteristics, ranging from the First Episode Social Functioning Scale (FESFS) with nine subscales to those who reported a few items (eg, part of a single subscale) of social functioning. Also, quality-of-life measures typically concentrated more on subjective evaluations of general life domains and were thus less focused on social functioning. The FESFS was the only measure to include an assessment of social activity on social medial; this is evaluated in a separate section below.

Psychometric Properties of the Measures

Out of all 58 included measures; 32 (55%) had previously been validated in patients with psychosis, 16 (28%) in a general psychiatric or clinical and community sample, 2 (3%) in a sample of patients with bipolar disorder, 2 (3%) in a sample of patients with depression, 2 (3%) in a sample of patients with somatic illness, 2 (3%) in a nonclinical sample, 1 (2%) in a sample of adolescents of parents with and without major depression, and 1 (2%) did not record any sample information. More data were available for reliability (53/58, 91%) than for validity (47/58, 81%). In general, lack of information prevented a comprehensive evaluation of the psychometric properties of most measure instruments. Theoretical foundation and construct validity was particularly poorly reported. When psychometric properties were reported, measures showed overall good validity and reliability regarding offline social functioning. The Social Functioning Scale, the Groningen Social Disability Schedule, and the Health of the Nation Outcomes Scale are examples of measure instruments with comprehensive reporting of this type of social functioning.

Measure Assessing Social Activity on Social Media

The FESFS was developed in 2014 [42] by the authors listing activities based on their experience with people with early psychosis and on reviews of existing measures of social functioning [40]. The FESFS is designed to measure social functioning in young people in the early stages of psychosis and was the only measure instrument identified in this review as addressing social activity on social media. The scale can be administered as observer-rated or self-report, with each item rating behavior—focus on frequency—and perceived individual ability. The FESFS comprises 34 items distributed on nine subscales assessing various domains of social functioning. The item language was intentionally constructed to fit the target group (eg, “hanging out with buddies” and “chatting on the net”). Two items, respectively from the items *Friends and activities* and *Living skills*, address social media activity explicitly: “I am really good in solo activities such as going to the gym, going to the movies, chatting on the net, taking lessons (music, painting, etc)” and “I am comfortable using the phone, Internet, or email to communicate.” The scale is cited five times, of which three of the cited articles include the measure developers as authors.

Scale validation was based on the self-report version. The validation sample included 203 people, with an average age of 24.5 years, diagnosed with a schizophrenia spectrum psychotic disorder, and with an average of 12.7 years of education. The nine factors showed good internal consistency, ranging from .63 to .80. Good convergent and discriminant validity, as well as good sensitivity to change, were also demonstrated. Three subscales had an inverse correlation with negative symptoms.

XSL•FO
RenderX

Discussion

Measures Should Include Contemporary Social Reality

Due to technological innovation and rapid alterations in the norms of social media usage, any instrument designed to measure social functioning, including social media activity, should encapsulate contemporary trends. The main finding of this review was that current measures of social functioning almost exclusively comprise offline social activity, with the sole exception of the FESFS, as discussed in a separate section below. This limitation is likely to reflect the time of development of currently used measures, as most were developed before the launch of the Internet in 1992, and only eight measures were developed or revised after the advent of Facebook in 2006. Many of these scales have good psychometric properties, which may be a good starting point if they were revised to include measures of social media activity. It is notable that the first measures developed were based on chronic inpatients (eg, the Interview Schedule for Social Interaction). However, there is now an emphasis on early intervention to target quality of life and younger early-stage patients, as opposed to chronic inpatients [109], and current measures fail to capture an important aspect of the current social context.

It is worth discussing whether the two most widely used categories of measures—social functioning and quality of life—are expedient. For instance, a number of the measures within these categories address social participation, while others address the more narrowly defined concept of social skills. In practice, then, choosing a measure from either category for evaluative purposes could potentially influence interpretations of findings.

Further, regarding validity, while the original psychometric assessment of some measures show good reliability and validity, they may lack ecological validity. For example, leaving out assessment of social media activity may lead to low scores on social functioning among young people with psychosis and, thus, increase the likelihood of false positives. Moreover, the core negative symptom of social withdrawal [8] may manifest differently in a social media context compared to an offline context. There is also a risk of social media addiction, negative social comparison, cyberbullying, as well as it being used to exclude real-life contacts [27,110], with potential negative consequences on illness course, outcome, and quality of life. Online social functioning measures should aim to be sensitive to these types of matters. Also, they should track symptom levels [2], change in social participation, and support that unfolds online [3]. In this regard, a survey found that adults with schizophrenia were as likely as adults without mental illness to form social relationships online, despite having fewer offline relationships, lower income, and less Internet access [4]. Compensating for symptoms that people with psychosis themselves experience that interfere with socializing in face-to-face encounters [111] may be a fruitful remedy for some of the obstacles associated with the enhanced levels of toxic loneliness and stigma associated with psychosis populations [5]. This type of information would also be important for treatment timing and tailoring.

Social Media Assessment

The FESFS represents an attempt to address contemporary forms of social functioning, including online activity. Additionally, the scale assesses both behavior and ability, which make a more nuanced assessment possible. However, the scale has fundamental limitations. The validation sample has an average age of 24.5 years, which is relatively high when aiming at early psychosis and UHR of psychosis. The subscales related to work and education are not satisfactorily validated, as only a small part of the validation sample was working or studying. Test-retest reliability for the scale has not been conducted and neither factorial structure nor construct validity has been confirmed. In addition, only the self-report version has been validated. Furthermore, the scale has only been cross-validated across context to a very limited extent [112]—as opposed to, for example, the Personal and Social Performance or the Psychosocial Functioning Scale—which implies uncertainty regarding robustness and usability. Further, the authors do not articulate a theoretical foundation for the scale, and scale content is derived from the scale authors' own listing of experience-based domains of social functioning.

Science and technology studies is a highly influential theoretical framework analyzing the entanglements of science, technology, and society [39]. A basic premise in *science and technology studies* is that technological innovation affects society and human behavior in fundamental ways. Specific technologies, such as social media, do not merely add to the possibilities of communication, but changes the nature of communication processes. Consequently, attempts to include technology-mediated communication processes should start from the premise that these probably do not reflect nontechnological communication. Compared with face-to-face contact, social media represents radically evolving platform structures and a more asynchronistic form of communication. However, it is unclear whether social media platforms require extra social flexibility or if they are adaptable to facilitate communication for persons who may have limitations in face-to-face social skills, such as the limitations typically found for individuals with active psychosis. It has been suggested that individuals with mental health problems may use social media to seek support. When compared to face-to-face interaction, social media allows more time for reflection before acting [113].

The FESFS "chatting on the net" item is defined as a solo activity and yet this may not reflect the experience of social media by individuals. Social media includes virtual communities allowing users to create a public profile, interact with real-life and virtual friends, and make new acquaintances. Social media engagements often seem to be a fundamental social activity [114]. Also, the FESFS defines using the Internet or email communication as a living skill. However, it is difficult to equate these technological skills as being representative of social activity or functioning. While the FESFS has been the first measure to attempt to capture social media activity, the measure requires significant further development for validity of measurement of contemporary social media engagements.

XSL•FO
RenderX

Future Research: Need for a Radical Change

The use of social media as a dimension of social functioning in psychosis is a complex issue and the knowledge base is limited. It is possible to explore social media behavior based on the most reliable and valid dimensions of currently available offline social functioning measures, such as the Social Functioning Scale. This scale provided the most comprehensive reporting of traditional psychometric properties for offline social functioning, including construct validity. With this scale, social skills or social behavior were registered as present or absent, thus removing the need for an evaluative decision. This could be a feasible starting point to track online social behavior. Some degree of social skill transfer between online and offline activities seems plausible. Additionally, it might be important to understand the relationship between more traditional measures of social skills and social media usage. In this regard, purely scientist-driven approaches have clear limitations. For example, the likely age gap between researchers and the target group of early psychosis, particularly the UHR segment, risks a lack of understanding of the social context. Therefore, a collaborative approach with the target group as codevelopers of the measure could remedy this shortcoming. The general omission of user involvement, which is highly prioritized and valued in most contemporary health care systems, is a major challenge to the validity of these measures [115]. We therefore propose a theoretical framework in which service users are involved, so as to explore social media as part of social functioning of young people with psychosis.

PNS was developed for interpreting and applying scientific results at the science-policy interface. PNS was tailored for situations where “facts [are] uncertain, values [are] in dispute, stakes [are] high, and decisions [are] urgent” [40]. The research field of social functioning in psychosis includes multiple theoretical perspectives, such as physiological, biological, evolutionary, social, and cultural perspectives. The complex nature of social functioning makes it difficult to indicate causality [1]. There are conflicts of interest causing tension between groups, such as the psychopharmaceutical industry, governments, professional associations, and user organizations [116-118]. The stakes are arguably high as social functioning impairment is regarded as a core symptom of serious mental illness, namely psychosis [119]. The PNS remedy is to communicate uncertainty, assess quality, and justify practice by including extended peer communities. In practical terms, the PNS framework ensures the inclusion of social components perceived as important by the target group. This will presumably lead to inclusion of new facets of social functioning that have been omitted by previous measures, and the risk of implementing outdated or ecologically invalid models is lowered.

Future reviews should take social media use or online activity into consideration when also evaluating social functioning measures in general patient populations. When developing and validating social functioning measures, researchers today should include social media activity: content, frequency, quality, and effects, both positive and negative.

Strengths and Limitations

The strengths of the study are evident in the study protocol being publicly available (ie, PROSPERO) before conducting the review, thus ensuring transparency, and the review was conducted according to PRISMA guidelines [41]. In addition, the inclusion of studies was determined by two independent raters and showed high interrater reliability.

The main purpose of this review was to assess to what degree social functioning measures included assessments of any online social activity. Hence, we applied broad inclusion criteria to avoid ignoring any potential measures. A side effect of this strategy was the inclusion of some measures that were not tailored to specifically target social functioning in general or psychosis specifically.

The conclusions drawn in this review may have been influenced by several of the included studies not reporting relevant psychometric properties. Although only one of the identified instruments specifically assessed social media activity, it cannot be ruled out that respondents may answer generic questions about social functioning with social media activity in mind. Another limitation is that each individual study was not assessed for key sources of biases (eg, sample characteristics). However, in line with previous research [20], it seems warranted to conclude that some studies were based on small samples and that most instruments were constructed and tested within Anglo-American cultures. Grey literature was not included. This will typically raise the risk of reporting bias, implying that the included studies represent selective research dissemination [120]. However, it should be emphasized that the aim was to identify instruments with a high level of use within the field and that the search was conducted in several literature databases. The included studies did use samples with somewhat different characteristics (eg, sex, age, and level of symptomatology), which may violate the transitivity assumption and, thus, questions direct comparisons across included studies.

Acknowledgments

A special thanks to the staff at the Medical Library of Stavanger University Hospital for assistance with the literature search.

Authors' Contributions

JB conceptualized and performed the current analyses and wrote the first draft. WTVH, HJS, SP had a special role in data collection, extraction, and analysis. SP also had a special role in ensuring that the chosen approach, PRISMA, was performed according to the guidelines. All authors were involved in study design, provided scientific oversight throughout the project, provided detailed comments about the paper across several drafts, and edited the paper.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Model search for replication.

[DOCX File, 15KB - jmir_v21i6e13957_app1.docx]

References

1. van Os J, Kenis G, Rutten BP. The environment and schizophrenia. Nature 2010 Nov 11;468(7321):203-212. [doi: 10.1038/nature09563] [Medline: 21068828]
2. Birnbaum M, Rizvi A, Choudhury MD, Ernala S, Cecchi G, Kane J. Identifying psychotic symptoms and predicting relapse trough social media. Schizophr Bull 2018 Apr 01;44(Suppl 1):S100. [doi: 10.1093/schbul/sby015.246]
3. Gowen K, Deschaine M, Gruttadara D, Markey D. Young adults with mental health conditions and social networking websites: Seeking tools to build community. Psychiatr Rehabil J 2012;35(3):245-250. [doi: 10.2975/35.3.2012.245.250] [Medline: 22246123]
4. Spinzy Y, Nitzan U, Becker G, Bloch Y, Fennig S. Does the Internet offer social opportunities for individuals with schizophrenia? A cross-sectional pilot study. Psychiatry Res 2012 Jul 30;198(2):319-320. [doi: 10.1016/j.psychres.2012.02.022] [Medline: 22440545]
5. Highton-Williamson E, Priebe S, Giacco D. Online social networking in people with psychosis: A systematic review. Int J Soc Psychiatry 2015 Feb 01;61(1):92-101. [doi: 10.1177/0020764014556392] [Medline: 25381145]
6. Stain H, Galletly C, Clark S, Wilson J, Killen EA, Anthes L, et al. Understanding the social costs of psychosis: The experience of adults affected by psychosis identified within the second Australian National Survey of Psychosis. Aust N Z J Psychiatry 2012 Sep;46(9):879-889. [doi: 10.1177/0004867412449060] [Medline: 22645395]
7. Kirkpatrick B, Fenton W, Carpenter W, Marder S. The NIMH-MATRICS consensus statement on negative symptoms. Schizophr Bull 2006 Apr;32(2):214-219 [FREE Full text] [doi: 10.1093/schbul/sbj053] [Medline: 16481659]
8. Marder S, Galderisi S. The current conceptualization of negative symptoms in schizophrenia. World Psychiatry 2017 Feb;16(1):14-24. [doi: 10.1002/wps.20385]
9. MacBeth A, Gumley A. Premorbid adjustment, symptom development and quality of life in first episode psychosis: A systematic review and critical reappraisal. Acta Psychiatr Scand 2008 Feb;117(2):85-99. [doi: 10.1111/j.1600-0447.2007.01134.x] [Medline: 18081922]
10. Zipursky RB. Why are the outcomes in patients with schizophrenia so poor? J Clin Psychiatry 2014;75(Suppl 2):20-24. [doi: 10.4088/JCP.13065su1.05]
11. Fusar-Poli P, Borgwardt S, Bechdolf A, Addington J, Riecher-Rössler A, Schultze-Lutter F, et al. The psychosis high-risk state: A comprehensive state-of-the-art review. JAMA Psychiatry 2013 Jan;70(1):107-120 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.269] [Medline: 23165428]
12. Fusar-Poli P, Rocchetti M, Sardella A, Avila A, Brandizzi M, Caverzasi E, et al. Disorder, not just state of risk: Meta-analysis of functioning and quality of life in people at high risk of psychosis. Br J Psychiatry 2015 Sep 2;207(3):198-206. [doi: 10.1192/bjp.bp.114.157115] [Medline: 26329563]
13. Fusar-Poli P, Smieskova R, Kempton M, Ho B, Andreasen N, Borgwardt S. Progressive brain changes in schizophrenia related to antipsychotic treatment? A meta-analysis of longitudinal MRI studies. Neurosci Biobehav Rev 2013 Sep;37(8):1680-1691 [FREE Full text] [doi: 10.1016/j.neubiorev.2013.06.001] [Medline: 23769814]
14. Bjornestad J, Hegelstad WTV, Joa I, Davidson L, Larsen TK, Melle I, et al. "With a little help from my friends": Social predictors of clinical recovery in first-episode psychosis. Psychiatry Res 2017 Dec;255:209-214. [doi: 10.1016/j.psychres.2017.05.041] [Medline: 28578180]
15. Turner D, McGlanaghy E, Cuijpers P, van der Gaag M, Karyotaki E, MacBeth A. A meta-analysis of social skills training and related interventions for psychosis. Schizophr Bull 2018 Dec 06;44(3):475-491 [FREE Full text] [doi: 10.1093/schbul/sbx146] [Medline: 29140460]
16. Lally J, Ajnakina O, Stubbs B, Cullinane M, Murphy KC, Gaughran F, et al. Remission and recovery from first-episode psychosis in adults: Systematic review and meta-analysis of long-term outcome studies. Br J Psychiatry 2017 Dec;211(6):350-358. [doi: 10.1192/bjp.bp.117.201475] [Medline: 28982659]
17. Thompson KN, McGorry PD, Harrigan SM. Recovery style and outcome in first-episode psychosis. Schizophr Res 2003 Jul;62(1-2):31-36 [FREE Full text] [doi: 10.1016/S0920-9964(02)00428-0]
18. Burns T, Patrick D. Social functioning as an outcome measure in schizophrenia studies. Acta Psychiatr Scand 2007 Dec;116(6):403-418 [FREE Full text] [doi: 10.1111/j.1600-0447.2007.01108.x] [Medline: 17941964]
19. Mueser KT, Tarrier N. Handbook Of Social Functioning In Schizophrenia. Boston, MA: Allyn & Bacon; 1998.
20. Priebe S. Social outcomes in schizophrenia. Br J Psychiatry 2018 Jan 02;191(S50):s15-s20. [doi: 10.1192/bjp.191.50.s15]

XSL•FO
RenderX

21. Duggan M. Pew Research Centre. Washington, DC: Pew Internet & American Life Project; 2015 Aug 19. Mobile messaging and social media 2015 URL: http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/ [accessed 2019-06-23]
22. Valmaggia L. The use of virtual reality in psychosis research and treatment. World Psychiatry 2017 Oct;16(3):246-247 [FREE Full text] [doi: 10.1002/wps.20443] [Medline: 28941114]
23. Villani M, Kovess-Masfety V. How do people experiencing schizophrenia spectrum disorders or other psychotic disorders use the Internet to get information on their mental health? Literature review and recommendations. JMIR Ment Health 2017 Jan 03;4(1):e1 [FREE Full text] [doi: 10.2196/mental.5946] [Medline: 28049620]
24. Statista. 2018 Nov 29. Number of social media users worldwide from 2010 to 2021 (in billions) URL: https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ [accessed 2019-06-23]
25. Perrin A. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2015 Oct 08. Social media usage: 2005-2015 URL: https://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/ [accessed 2019-06-23]
26. Lenhart A, Duggan M, Perrin A, Stepler R, Rainie L, Parker K. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2015 Apr 09. Teens, social media, & technology overview 2015: Smartphones facilitate shifts in communication landscape for teens URL: http://www.pewresearch.org/wp-content/uploads/sites/9/2015/04/PI_TeensandTech_Update2015_0409151.pdf [accessed 2019-06-23]
27. Kaplan A, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Bus Horiz 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
28. Torous J, Keshavan M. The role of social media in schizophrenia: Evaluating risks, benefits, and potential. Curr Opin Psychiatry 2016 May;29(3):190-195. [doi: 10.1097/YCO.0000000000000246] [Medline: 26967314]
29. Alvarez-Jimenez M, Alcazar-Corcoles M, González-Blanch C, Bendall S, McGorry P, Gleeson J. Online, social media and mobile technologies for psychosis treatment: A systematic review on novel user-led interventions. Schizophr Res 2014 Jun;156(1):96-106. [doi: 10.1016/j.schres.2014.03.021] [Medline: 24746468]
30. Firth J, Torous J. Smartphone apps for schizophrenia: A systematic review. JMIR Mhealth Uhealth 2015 Nov 06;3(4):e102 [FREE Full text] [doi: 10.2196/mhealth.4930] [Medline: 26546039]
31. Naslund J, Aschbrenner K, McHugo G, Unützer J, Marsch L, Bartels S. Exploring opportunities to support mental health care using social media: A survey of social media users with mental illness. Early Interv Psychiatry 2019 Jun;13(3):405-413. [doi: 10.1111/eip.12496] [Medline: 29052947]
32. Batra S, Baker R, Wang T, Forma F, DiBiasi F, Peters-Strickland T. Digital health technology for use in patients with serious mental illness: A systematic review of the literature. Med Devices (Auckl) 2017;10:237-251 [FREE Full text] [doi: 10.2147/MDER.S144158] [Medline: 29042823]
33. Kumar D, Tully L, Iosif A, Zakskorn LN, Nye KE, Zia A, et al. A mobile health platform for clinical monitoring in early psychosis: Implementation in community-based outpatient early psychosis care. JMIR Ment Health 2018 Feb 27;5(1):e15 [FREE Full text] [doi: 10.2196/mental.8551] [Medline: 29487044]
34. Naslund J, Aschbrenner K, Marsch L, McHugo G, Bartels S. Facebook for supporting a lifestyle intervention for people with major depressive disorder, bipolar disorder, and schizophrenia: An exploratory study. Psychiatr Q 2018 Dec;89(1):81-94 [FREE Full text] [doi: 10.1007/s11126-017-9512-0] [Medline: 28470468]
35. Mueller N, Panch T, Macias C, Cohen B, Ongur D, Baker J. Using smartphone apps to promote psychiatric rehabilitation in a peer-led community support program: Pilot study. JMIR Ment Health 2018 Aug 15;5(3):e10092 [FREE Full text] [doi: 10.2196/10092] [Medline: 30111526]
36. Ladea M, Bran M, Marcel Claudiu S. Online destigmatization of schizophrenia: A Romanian experience. Eur Psychiatry 2016 Mar;33:S276. [doi: 10.1016/j.eurpsy.2016.01.736]
37. Alverez M, Thomas N. Innovations in technology-based treatments and management for psychosis: Moderated online social therapy (Horyzons) and self-management and recovery technology. In: Proceedings of the Royal Australian and New Zealand College of Psychiatrists Congress (RANZCP 2017). 2017 Presented at: The Royal Australian and New Zealand College of Psychiatrists Congress (RANZCP 2017); April 30-May 4, 2017; Adelaide, Australia.
38. Alvarez-Jimenez M, Gleeson J, Bendall S, Rice S, D'Alfonso S, Eleftheriadis D, et al. Momentum: A novel online social media, mindfulness, and strengths-based intervention to promote functional recovery in ultra high-risk (UHR) patients. Schizophr Bull 2017 Mar;43(Suppl 1):S206-S207. [doi: 10.1093/schbul/sbx024.121]
39. Jasanoff S. The Ethics of Invention: Technology and the Human Future. New York, NY: WW Norton & Company; 2016.
40. Funtowicz SO, Ravetz JR. Science for the post-normal age. Futures 1993 Sep;25(7):739-755. [doi: 10.1016/0016-3287(93)90022-L]
41. Hutton B, Salanti G, Caldwell D, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. Ann Intern Med 2015 Jun 02;162(11):777-784. [doi: 10.7326/M14-2385] [Medline: 26030634]
42. Lecomte T, Corbière M, Ehmann T, Addington J, Abdel-Baki A, Macewan B. Development and preliminary validation of the First Episode Social Functioning Scale for early psychosis. Psychiatry Res 2014 May 30;216(3):412-417. [doi: 10.1016/j.psychres.2014.01.044] [Medline: 24613006]

XSL•FO
RenderX

43. Forsyth K, Lai J, Kielhofner G. The Assessment of Communication and Interaction Skills (ACIS): Measurement properties. Br J Occup Ther 2016 Nov 05;62(2):69-74. [doi: 10.1177/030802269906200208]
44. Hsu W, Pan AW, Chen TJ. A psychometric study of the Chinese version of the assessment of communication and interaction skills. Occup Ther Health Care 2008;22(2-3):177-185. [doi: 10.1080/07380570801991818] [Medline: 23941383]
45. Eisen S, Wilcox M, Leff H, Schaefer E, Culhane M. Assessing behavioral health outcomes in outpatient programs: Reliability and validity of the BASIS-32. J Behav Health Serv Res 1999 Feb;26(1):5-17. [doi: 10.1007/BF02287790]
46. Wallace C, Lecomte T, Wilde J, Liberman R. CASIG: A consumer-centered assessment for planning individualized treatment and evaluating program outcomes. Schizophr Res 2001 May;50(1-2):105-119. [doi: 10.1016/S0920-9964(00)00068-2]
47. Brekke JS, Levin S, Wolkon GH, Sobel E, Slade E. Psychosocial functioning and subjective experience in schizophrenia. Schizophr Bull 1993 Jan 01;19(3):599-608. [doi: 10.1093/schbul/19.3.599] [Medline: 8235461]
48. Üstün T, Kostanjsek N, Chatterji S, editors. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 2.0). Geneva, Switzerland: World Health Organization; 2010.
49. Rosa A, Sánchez-Moreno J, Martínez-Aran A, Salamero M, Torrent C, Reinares M, et al. Validity and reliability of the Functioning Assessment Short Test (FAST) in bipolar disorder. Clin Pract Epidemiol Ment Health 2007 Jun 07;3:5 [FREE Full text] [doi: 10.1186/1745-0179-3-5] [Medline: 17555558]
50. Llorca P, Lançon C, Lancrenon S, Bayle FJ, Caci H, Rouillon F, et al. The "Functional Remission of General Schizophrenia" (FROGS) scale: Development and validation of a new questionnaire. Schizophr Res 2009 Sep;113(2-3):218-225. [doi: 10.1016/j.schres.2009.04.029] [Medline: 19464855]
51. Dausch BM, Miklowitz DJ, Richards JA. Global Assessment of Relational Functioning Scale (GARF): II. Reliability and validity in a sample of families of bipolar patients. Fam Process 1996 Jun;35(2):175-189. [doi: 10.1111/j.1545-5300.1996.00175.x]
52. Cornblatt B, Auther A, Niendam T, Smith CW, Zinberg J, Bearden CE, et al. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. Schizophr Bull 2007 May;33(3):688-702 [FREE Full text] [doi: 10.1093/schbul/sbm029] [Medline: 17440198]
53. Schützwohl M, Jarosz-Nowak J, Briscoe J, Szajowski K, Kallert T. Inter-rater reliability of the Brief Psychiatric Rating Scale and the Groningen Social Disabilities Schedule in a European multi-site randomized controlled trial on the effectiveness of acute psychiatric day hospitals. Int J Methods Psychiatr Res 2006 Nov;12(4):197-207. [doi: 10.1002/mpr.157]
54. Van der Does AJ, Linszen DH, Dingemans PM, Nugter MA, Scholte WF. A dimensional and categorical approach to the symptomatology of recent-onset schizophrenia. J Nerv Ment Dis 1993 Dec;181(12):744-749. [Medline: 8254326]
55. Orrell M, Yard P, Handysides J, Schapira R. Validity and reliability of the Health of the Nation Outcome Scales in psychiatric patients in the community. Br J Psychiatry 1999 May;174:409-412. [Medline: 10616606]
56. McConkey R, Walsh J. An index of social competence for use in determining the service needs of mentally handicapped adults. J Ment Defic Res 1982 Mar;26(Pt 1):47-61. [Medline: 7077654]
57. Horowitz L, Rosenberg S, Baer B, Ureño G, Villaseñor VS. Inventory of interpersonal problems: Psychometric properties and clinical applications. J Consult Clin Psychol 1988 Dec;56(6):885-892. [Medline: 3204198]
58. Henderson S, Duncan-Jones P, Byrne D, Scott R. Measuring social relationships. The Interview Schedule for Social Interaction. Psychol Med 1980 Nov;10(4):723-734. [Medline: 7208730]
59. Parker G, Rosen A, Emdur N, Hadzi-Pavlov D. The Life Skills Profile: Psychometric properties of a measure assessing function and disability in schizophrenia. Acta Psychiatr Scand 1991 Feb;83(2):145-152. [doi: 10.1111/j.1600-0447.1991.tb07381.x]
60. Barker S, Barron N, McFarland B, Bigelow D. A community ability scale for chronically mentally Ill consumers: Part I. Reliability and validity. Community Ment Health J 1994 Aug;30(4):363-383. [doi: 10.1007/BF02207489]
61. Patrick D, Burns T, Morosini P, Rothman M, Gagnon DD, Wild D, et al. Reliability, validity and ability to detect change of the clinician-rated Personal and Social Performance scale in patients with acute symptoms of schizophrenia. Curr Med Res Opin 2009 Jan 07;25(2):325-338. [doi: 10.1185/03007990802611919]
62. Huprich SK, Sanford K, Smith M. Psychometric evaluation of the depressive personality disorder inventory. J Pers Disord 2002 Jun;16(3):255-269. [Medline: 12136681]
63. Valencia M, Rojas E, González C, Ramos L, Villatoro J. Evaluation of social functioning in patients at a health center [Article in Spanish]. Salud Publica Mex 1989;31(5):674-687. [Medline: 2609230]
64. Corrigan P, Giffort D, Rashid F, Leary M, Okeke I. Recovery as a psychological construct. Community Ment Health J 1999 Jun;35(3):231-239. [Medline: 10401893]
65. Foreman EI, Baker R. A validity study of a new rating scale for psychiatric patients. Acta Psychiatr Scand 1986 Jan;73(1):101-108. [doi: 10.1111/j.1600-0447.1986.tb02674.x]
66. Goodman S, Sewell D, Cooley E, Leavitt N. Assessing levels of adaptive functioning: The Role Functioning Scale. Community Ment Health J 1993 Apr;29(2):119-131. [Medline: 8500285]
67. Padmavathi R, Thara R, Srinivasan L, Kumar S. Scarf social functioning index. Indian J Psychiatry 1995 Oct;37(4):161-164 [FREE Full text] [Medline: 21743742]

XSL•FO
RenderX

68. Barrowclough C, Tarrier N, Humphreys L, Ward J, Gregg L, Andrews B. Self-esteem in schizophrenia: Relationships between self-evaluation, family attitudes, and symptomatology. J Abnorm Psychol 2003 Feb;112(1):92-99. [Medline: 12653417]
69. Russo J, Trujillo C, Wingerson D, Decker K, Ries R, Wetzler H, et al. The MOS 36-Item Short Form Health Survey: Reliability, validity, and preliminary findings in schizophrenic outpatients. Med Care 1998 May;36(5):752-756. [Medline: 9596066]
70. Bosc M, Dubini A, Polin V. Development and validation of a social functioning scale, the Social Adaptation Self-evaluation Scale. Eur Neuropsychopharmacol 1997 Apr;7 Suppl 1:S57-S70; discussion S71. [Medline: 9169311]
71. Bech P, Lunde M, Undén M. Social Adaptation Self-evaluation Scale (SASS): Psychometric analysis as outcome measure in the treatment of patients with major depression in the remission phase. Int J Psychiatry Clin Pract 2002;6(3):141-146. [doi: 10.1080/136515002760276063] [Medline: 24945200]
72. Harvey P, Davidson M, Mueser K, Parrella M, White L, Powchik P. Social-Adaptive Functioning Evaluation (SAFE): A rating scale for geriatric psychiatric patients. Schizophr Bull 1997;23(1):131-145. [doi: 10.1093/schbul/23.1.131] [Medline: 9050119]
73. John K, Gammon D, Prusoff B, Warner V. The Social Adjustment Inventory for Children and Adolescents (SAICA): Testing of a new semistructured interview. J Am Acad Child Adolesc Psychiatry 1987 Nov;26(6):898-911. [Medline: 3429410]
74. Weissman MM. Assessment of social adjustment by patient self-report. Arch Gen Psychiatry 1976 Sep 01;33(9):1111. [doi: 10.1001/archpsyc.1976.01770090101010]
75. Wykes T, Wing J. A ward in a house: Accommodation for "new" long-stay patients. Acta Psychiatr Scand 1982 May;65(5):315-330. [Medline: 6810654]
76. Wykes T, Sturt E. The measurement of social behaviour in psychiatric patients: An assessment of the reliability and validity of the SBS schedule. Br J Psychiatry 1986 Jan;148:1-11. [Medline: 3082403]
77. Tyrer P, Nur U, Crawford M, Karlsen S, McLean C, Rao B, et al. The Social Functioning Questionnaire: A rapid and robust measure of perceived functioning. Int J Soc Psychiatry 2005 Sep;51(3):265-275. [Medline: 16252794]
78. Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S. The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. Br J Psychiatry 1990 Dec;157:853-859. [Medline: 2289094]
79. Kawata A, Revicki DA. Reliability and validity of the social integration survey (SIS) in patients with schizophrenia. Qual Life Res 2008 Feb;17(1):123-135. [doi: 10.1007/s11136-007-9288-z] [Medline: 18058036]
80. Roy-Byrne P, Dagadakis C, Unutzer J, Ries R. Evidence for limited validity of the revised global assessment of functioning scale. Psychiatr Serv 1996 Aug;47(8):864-866. [doi: 10.1176/ps.47.8.864] [Medline: 8837160]
81. Saraswat N, Rao K, Subbakrishna D, Gangadhar B. The Social Occupational Functioning Scale (SOFS): A brief measure of functional status in persons with schizophrenia. Schizophr Res 2006 Jan 31;81(2-3):301-309. [doi: 10.1016/j.schres.2005.09.008] [Medline: 16256309]
82. McPheeters H. Statewide mental health outcome evaluation: A perspective of two southern states. Community Ment Health J 1984;20(1):44-55. [Medline: 6723258]
83. Schneider L, Struening E. SLOF: A behavioral rating scale for assessing the mentally ill. Soc Work Res Abstr 1983;19(3):9-21. [Medline: 10264257]
84. Jolley S, Garety P, Ellett L, Kuipers E, Freeman D, Bebbington PE, et al. A validation of a new measure of activity in psychosis. Schizophr Res 2006 Jul;85(1-3):288-295. [doi: 10.1016/j.schres.2006.03.012] [Medline: 16626943]
85. Jolley S, Garety P, Dunn G, White J, Aitken M, Challacombe F, et al. A pilot validation study of a new measure of activity in psychosis. Soc Psychiatry Psychiatr Epidemiol 2005 Nov;40(11):905-911. [doi: 10.1007/s00127-005-0982-x] [Medline: 16220212]
86. Hawthorne G. Assessing utility where short measures are required: Development of the short Assessment of Quality of Life-8 (AQoL-8) instrument. Value Health 2009 Sep;12(6):948-957 [FREE Full text] [doi: 10.1111/j.1524-4733.2009.00526.x] [Medline: 19490558]
87. Lam C, van Weel C, Lauder IJ. Can the Dartmouth COOP/WONCA charts be used to assess the functional status of Chinese patients? Fam Pract 1994 Mar;11(1):85-94. [doi: 10.1093/fampra/11.1.85] [Medline: 8034158]
88. Nelson E, Wasson J, Kirk J, Keller A, Clark D, Dietrich A, et al. Assessment of function in routine clinical practice: Description of the COOP Chart method and preliminary findings. J Chronic Dis 1987;40 Suppl 1:55S-69S. [Medline: 3597698]
89. van Nieuwenhuizen C, Schene A, Koeter M, Huxley P. The Lancashire Quality of Life Profile: Modification and psychometric evaluation. Soc Psychiatry Psychiatr Epidemiol 2001 Jan;36(1):36-44. [Medline: 11320806]
90. Priebe S, Huxley P, Knight S, Evans S. Application and results of the Manchester Short Assessment of Quality of Life (MANSA). Int J Soc Psychiatry 1999;45(1):7-12. [doi: 10.1177/002076409904500102] [Medline: 10443245]
91. Björkman T, Svensson B. Quality of life in people with severe mental illness. Reliability and validity of the Manchester Short Assessment of Quality of Life (MANSA). Nord J Psychiatry 2005;59(4):302-306. [doi: 10.1080/08039480500213733] [Medline: 16195135]

XSL•FO
RenderX

92. Pukrop R, Möller HJ, Steinmeyer E. Quality of life in psychiatry: A systematic contribution to construct validation and the development of the integrative assessment tool "modular system for quality of life". Eur Arch Psychiatry Clin Neurosci 2000;250(3):120-132. [Medline: 10941986]
93. Bergner M, Bobbitt R, Carter W, Gilson B. The Sickness Impact Profile: Development and final revision of a health status measure. Med Care 1981 Aug;19(8):787-805. [Medline: 7278416]
94. Gerety M, Cornell J, Mulrow C, Tuley M, Hazuda HP, Lichtenstein M, et al. The Sickness Impact Profile for nursing homes (SIP-NH). J Gerontol 1994 Jan;49(1):M2-M8. [Medline: 8282976]
95. Endicott J, Nee J, Harrison W, Blumenthal R. Quality of Life Enjoyment and Satisfaction Questionnaire: A new measure. Psychopharmacol Bull 1993;29(2):321-326. [Medline: 8290681]
96. Ritsner MS, Awad AG, editors. Quality of Life Impairment in Schizophrenia, Mood and Anxiety Disorders: New Perspectives on Research and Treatment. Dordrecht, the Netherlands: Springer; 2007.
97. Swan A, Watson H, Nathan P. Quality of life in depression: An important outcome measure in an outpatient cognitive-behavioural therapy group programme? Clin Psychol Psychother 2009;16(6):485-496. [doi: 10.1002/cpp.588] [Medline: 19475699]
98. Ferrans CE, Powers MJ. Psychometric assessment of the quality of life index. Res Nurs Health 1992 Feb;15(1):29-38. [doi: 10.1002/nur.4770150106]
99. Lehman A. Measures of quality of life among persons with severe and persistent mental disorders. Soc Psychiatry Psychiatr Epidemiol 1996 Mar;31(2):78-88. [Medline: 8881088]
100. Frisch MB, Cornell J, Villanueva M, Retzlaff PJ. Clinical validation of the Quality of Life Inventory. A measure of life satisfaction for use in treatment planning and outcome assessment. Psychol Assess 1992;4(1):92-101. [doi: 10.1037/1040-3590.4.1.92]
101. Heinrichs DW, Hanlon TE, Carpenter WT. The Quality of Life Scale: An instrument for rating the schizophrenic deficit syndrome. Schizophr Bull 1984 Jan 01;10(3):388-398. [doi: 10.1093/schbul/10.3.388] [Medline: 6474101]
102. Kaplan R, Atkins C, Timms R. Validity of a quality of well-being scale as an outcome measure in chronic obstructive pulmonary disease. J Chronic Dis 1984;37(2):85-95. [Medline: 6420431]
103. Kaplan RM, Anderson JP, Ganiats TG. The Quality of Well-being Scale: Rationale for a single quality of life index. In: Walker SR, Rosser RM, editors. Quality of Life Assessment: Key Issues in the 1990s. Dordrecht, the Netherlands: Springer; 1993:65-94.
104. Test M, Greenberg J, Long J, Brekke J, Burke S. Construct validity of a measure of subjective satisfaction with life of adults with serious mental illness. Psychiatr Serv 2005 Mar;56(3):292-300. [doi: 10.1176/appi.ps.56.3.292] [Medline: 15746503]
105. Boyer L, Simeoni M, Loundou A, D'Amato T, Reine G, Lancon C, et al. The development of the S-QoL 18: A shortened quality of life questionnaire for patients with schizophrenia. Schizophr Res 2010 Aug;121(1-3):241-250. [doi: 10.1016/j.schres.2010.05.019] [Medline: 20541912]
106. Wilkinson G, Hesdon B, Wild D, Cookson R, Farina C, Sharma V, et al. Self-report quality of life measure for people with schizophrenia: The SQLS. Br J Psychiatry 2000 Jul;177:42-46. [Medline: 10945087]
107. Becker M, Diamond R, Sainfort F. A new patient focused index for measuring quality of life in persons with severe and persistent mental illness. Qual Life Res 1993 Aug;2(4):239-251. [Medline: 8220359]
108. World Health Organization. The World Health Organization Quality of Life assessment (WHOQOL): Position paper from the World Health Organization. Soc Sci Med 1995 Nov;41(10):1403-1409. [Medline: 8560308]
109. Fusar-Poli P, McGorry P, Kane J. Improving outcomes of first-episode psychosis: An overview. World Psychiatry 2017 Oct;16(3):251-265 [FREE Full text] [doi: 10.1002/wps.20446] [Medline: 28941089]
110. Andreassen CS, Pallesen S. Social network site addiction: An overview. Curr Pharm Des 2014;20(25):4053-4061. [Medline: 24001298]
111. Dickerson F, Sommerville J, Origoni AE, Ringel NB, Parente F. Outpatients with schizophrenia and bipolar I disorder: Do they differ in their cognitive and social functioning? Psychiatry Res 2001 May 10;102(1):21-27. [Medline: 11368836]
112. Wang Y, Yeh YH, Tsang SM, Liu WH, Shi HS, Li Z, et al. Social functioning in Chinese college students with and without schizotypal personality traits: An exploratory study of the Chinese version of the First Episode Social Functioning Scale. PLoS One 2013;8(5):e61115 [FREE Full text] [doi: 10.1371/journal.pone.0061115] [Medline: 23690922]
113. Andreassen C, Torsheim T, Brunborg G, Pallesen S. Development of a Facebook Addiction Scale. Psychol Rep 2012 Apr;110(2):501-517. [doi: 10.2466/02.09.18.PR0.110.2.501-517] [Medline: 22662404]
114. Kuss D, Griffiths M. Online social networking and addiction: A review of the psychological literature. Int J Environ Res Public Health 2011 Dec;8(9):3528-3552 [FREE Full text] [doi: 10.3390/ijerph8093528] [Medline: 22016701]
115. Wykes T. Great expectations for participatory research: What have we achieved in the last ten years? World Psychiatry 2014 Feb;13(1):24-27 [FREE Full text] [doi: 10.1002/wps.20086] [Medline: 24497239]
116. Davidson L, Schmutte T, Dinzeo T, Andres-Hyman R. Remission and recovery in schizophrenia: Practitioner and patient perspectives. Schizophr Bull 2008 Jan;34(1):5-8 [FREE Full text] [doi: 10.1093/schbul/sbm122] [Medline: 17984297]

XSL•FO
RenderX

117. Slade M, Leamy M, Bacon F, Janosik M, Le Boutillier C, Williams J, et al. International differences in understanding recovery: Systematic review. Epidemiol Psychiatr Sci 2012 Dec;21(4):353-364. [doi: 10.1017/S2045796012000133] [Medline: 22794507]
118. Sohler N, Adams B, Barnes D, Cohen G, Prins S, Schwartz S. Weighing the evidence for harm from long-term treatment with antipsychotic medications: A systematic review. Am J Orthopsychiatry 2016;86(5):477-485 [FREE Full text] [doi: 10.1037/ort0000106] [Medline: 26652608]
119. Evensen S, Wisløff T, Lystad J, Bull H, Ueland T, Falkum E. Prevalence, employment rate, and cost of schizophrenia in a high-income welfare society: A population-based study using comprehensive health and welfare registers. Schizophr Bull 2016 Mar;42(2):476-483 [FREE Full text] [doi: 10.1093/schbul/sbv141] [Medline: 26433216]
120. Kirkham J, Altman D, Williamson P. Bias due to changes in specified outcomes during the systematic review process. PLoS One 2010 Mar 22;5(3):e9810 [FREE Full text] [doi: 10.1371/journal.pone.0009810] [Medline: 20339557]

Abbreviations

DARE: Database of Abstracts of Reviews of Effects
DAS: Disability Assessment Schedule
DAS-II-sv: Disability Assessment Schedule—II: Schizophrenia Version
DSM: Diagnostic and Statistical Manual of Mental Disorders
FESFS: First Episode Social Functioning Scale
ICD: International Classification of Diseases
MeSH: Medical Subject Headings
PNS: post-normal science
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
SDSS: Social Disability Screening Schedule
UHR: ultrahigh risk
WHO: World Health Organization
WHO-DAS: World Health Organization Disability Assessment Schedule
WHOQOL-BREF: World Health Organization Quality of Life Brief Version

Edited by G Eysenbach; submitted 08.03.19; peer-reviewed by O Ness, F Dark, G Snethen; comments to author 03.04.19; revised version received 06.05.19; accepted 10.05.19; published 28.06.19.

Please cite as:
Bjornestad J, Hegelstad WTV, Berg H, Davidson L, Joa I, Johannessen JO, Melle I, Stain HJ, Pallesen S
Social Media and Social Functioning in Psychosis: A Systematic Review
J Med Internet Res 2019;21(6):e13957
URL: http://www.jmir.org/2019/6/e13957/
doi:10.2196/13957
PMID:31254338

©Jone Bjornestad, Wenche Ten Velden Hegelstad, Henrik Berg, Larry Davidson, Inge Joa, Jan Olav Johannessen, Ingrid Melle, Helen J Stain, Ståle Pallesen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Unintended Consequences of Nationwide Electronic Health Record Adoption: Challenges and Opportunities in the Post-Meaningful Use Era

Tiago K Colicchio[1], PhD, MBA; James J Cimino[1], MD; Guilherme Del Fiol[2], MD, PhD

[1]Informatics Institute, University of Alabama at Birmingham, Birmingham, AL, United States

[2]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Tiago K Colicchio, PhD, MBA
Informatics Institute
University of Alabama at Birmingham
1900 University Boulevard
Birmingham, AL, 35294
United States
Phone: 1 2059960735
Email: tcolicchio@uabmc.edu

Abstract

The US health system has recently achieved widespread adoption of electronic health record (EHR) systems, primarily driven by financial incentives provided by the Meaningful Use (MU) program. Although successful in promoting EHR adoption and use, the program, and other contributing factors, also produced important unintended consequences (UCs) with far-reaching implications for the US health system. Based on our own experiences from large health information technology (HIT) adoption projects and a collection of key studies in HIT evaluation, we discuss the most prominent UCs of MU: failed expectations, EHR market saturation, innovation vacuum, physician burnout, and data obfuscation. We identify challenges resulting from these UCs and provide recommendations for future research to empower the broader medical and informatics communities to realize the full potential of a now digitized health system. We believe that fixing these unanticipated effects will demand efforts from diverse players such as health care providers, administrators, HIT vendors, policy makers, informatics researchers, funding agencies, and outside developers; promotion of new business models; collaboration between academic medical centers and informatics research departments; and improved methods for evaluations of HIT.

(J Med Internet Res 2019;21(6):e13313) doi:10.2196/13313

KEYWORDS

meaningful use; medical informatics applications; adoption

Introduction

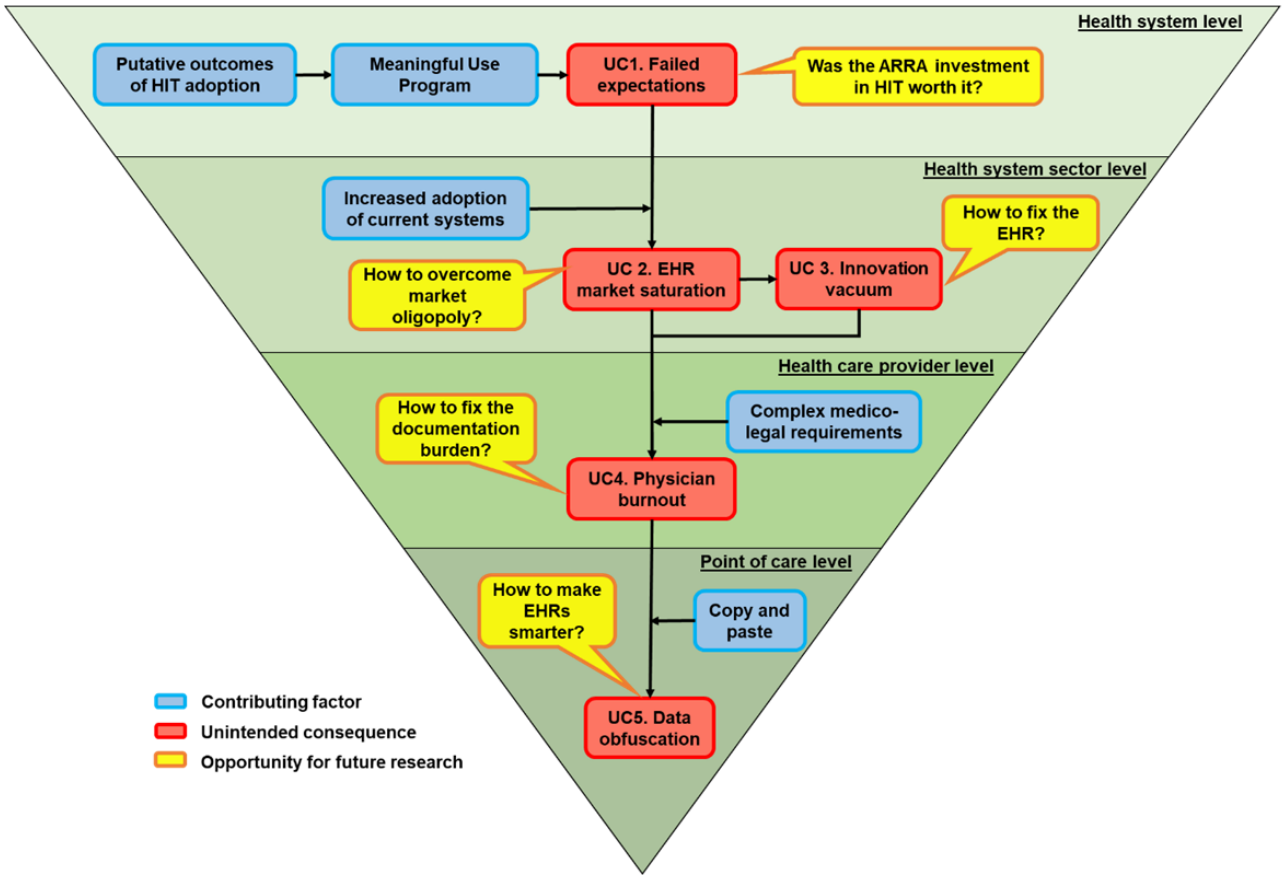
When humans created the cities to enable surplus food, labor division, and trade, the city itself generated new modalities of problems such as disease and violence. The American sociologist Robert K. Merton (1910-2013) coined the term unintended consequences (UCs) to describe these antagonistic elements inherent in any human endeavor [1]. The health care industry, which in the United States has reached near universal adoption of electronic health record (EHR) systems, is no exception.

Calls for nationwide adoption of EHRs [2] finally came to fruition when the US Congress passed the Health Information Technology for Economic and Clinical Health (HITECH) Act into law in 2009 [3], establishing the Meaningful Use (MU) program. As a result of MU, EHR adoption among US hospitals increased an impressive 8-fold in 6 years, and today, 9 in 10 hospitals use a government-certified EHR, and adoption among office-based physicians is above 80% [4]. However, although successful in promoting its intended consequences (EHR adoption and use), the program, and other contributing factors, also produced important UCs, with effects that range from the health system level all the way to the point of care level. Many recent publications have criticized MU and particularly EHRs; however, little attention has been dedicated to promoting effective solutions. Although previous articles have elicited emerging health information technology (HIT) UCs such as decreased patient-provider interaction, security breaches, and overdependence on technology [5] and proposed a research

XSL•FO
RenderX

agenda to fixing the EHR [6], such reports were produced during the MU implementation, and therefore, their conclusions were made before the US health system had been exposed to the effects of nationwide EHR adoption. On the basis of our own experiences from large-scale HIT adoption projects and a collection of key studies in HIT evaluation, we discuss the most prominent UCs of MU (Figure 1) and provide recommendations for future research to empower the broader medical and informatics communities to realize the full potential of a now digitized health system.

Figure 1. Unintended consequences of Meaningful Use, their contributing factors, and opportunities for future research from the broadest to the most specific level. ARRA: American Recovery and Reinvestment Act; EHR: electronic health record; HIT: health information technology; UC: unintended consequence.



Unintended Consequence 1: Failed Expectations

Recent systematic reviews have found that most HIT evaluations published before MU reported predominantly positive outcomes [7,8]. These outcomes served as the foundation for the MU program and have produced a hype around HIT. Such a hype led to a nationwide adoption of commercial EHRs with high expectations for improving the US health care cost and quality [9]. However, after 4 years of nationwide EHR adoption, health care in the United States is still the most expensive and lags behind in some quality outcomes when compared with other developed countries [10], which indicates that the expected benefits of a digital health system have not yet materialized [11-14]. As the adoption of commercial EHRs increased, new, unanticipated modalities of problems emerged [5]. The first systematic review of HIT impact published after MU continued to find mostly positive results; however, it also reported that 19% of the studies found no significant HIT impact, and the lack of negative outcomes is likely explained by publication bias [15].

The same systematic reviews that have reported positive findings have also reported several mixed results, which leaves unanswered questions as to the impact of HIT on quality, productivity, and safety. Furthermore, studies from other industries demonstrate that IT adoption rarely produces positive results if not accompanied by complementary factors or investments [16]. Several internal and external factors have been identified as potentially affecting care outcomes during HIT interventions [17], which suggests that previous studies may have been subjected to similar context-dependent factors, as they are common to HIT interventions [18,19]. Pre-MU studies are being criticized for relying on weak research designs such as short-term pretest-posttests and for the use of a small set of nonconsensus measurements [8,12,20]. The latter is an important barrier to the reproducibility of studies [21] and to the comparison of outcomes across studies [20], which prevents more comprehensive assessments of HIT impact and produces questions regarding the strength of the evidence supporting HIT effectiveness [22]. The lack of consistent evidence resulting from the use of poorly designed studies indicates that what others have called *positive outcomes* [7,8] are in fact *putative outcomes*. It has been estimated that without improved research

XSL•FO
RenderX

methods, around 100 hypotheses per year will continue to be tested without providing any valuable knowledge [23].

With insufficient evidence to support the hype around HIT and generalizable effects of HIT across care outcomes, settings, and EHR systems, an important question remains unanswered: was the over 20-billion-dollar investment in HIT from the America Recovery and Reinvestment Act (ARRA) worth it?

The Path Forward

Implementation of a new EHR will inevitably add to the complexity of the several aspects of care, and as users adapt to the system, they demand new customizations [24]. These customizations are often added to updated EHR versions that demand extensive local testing and an implementation process almost as complex, risky, and labor intensive as the implementation of a newly adopted EHR. In such a scenario, simple pretest-posttest designs are ineffective [25]. A paradigm shift on the choice of research designs for HIT studies is needed to produce more longitudinal evaluations able to detect time-sensitive effects common to HIT interventions [26] and to assess a large set of measures capable of detecting the diverse effects of such interventions [11,12]. Furthermore, as HIT interventions are subject to context-dependent factors, assessment of potential covariates is of paramount importance, as demonstrated elsewhere [17]. A better understanding of the full impact of HIT on the US health system will demand more comprehensive evaluations that assess a large sample of agreed-upon measures shared across researchers to allow comparison of outcomes across studies by future systematic reviews—and potential meta-analyses. In addition to increasing our understating of HIT impact on a national scale, such an approach has the potential to produce compelling evidence to the need for improving HIT effectiveness and can lead us to a more realistic assessment of the real value of the ARRA investment in HIT.

Unintended Consequence 2: Electronic Health Record Market Saturation

The time frame to implement MU's certification criteria was constrained, and the larger EHR vendors more rapidly complied with the criteria, contributing to an increased adoption of systems with established market share [27]. In 2017, the top 3 US HIT vendors shared 66% of the EHR market for acute care hospitals, which includes most large academic medical centers [28,29]. Given the complexity and high cost involved in implementing a commercial EHR, health care organizations are unlikely to change an EHR vendor anytime soon, causing a saturation of the US EHR market.

The Path Forward

As new, expensive EHR implementations become rarer, EHR vendors will be forced to find new business models to remain profitable. This path is evolving through initiatives such as the Substitutable Medical Applications & Reusable Technologies (SMART), which coupled with data standards, such as Fast Healthcare Interoperability Resources (FHIR), is enabling development of third-party applications seamlessly connected to commercial EHRs. Such applications have the potential to replace or augment commercial EHRs' functionality, in a model similar to the mobile phone industry [30]. To providers, such an approach represents an interesting opportunity to expand, customize, or replace EHR functionality as needed; to EHR vendors, it represents an opportunity to diversify their products, solutions, and sources of income. However, the saturation of the national market has produced a situation analogous to an oligopoly, and the path to producing new business models is unclear. Although some vendors seem to be open to the idea of having external applications connected to their EHR, others intend to charge providers per FHIR transaction, which will eventually hamper use of external applications. In addition, the 2 leading US EHR vendors are increasing their global presence [31], which may help to keep them financially sustainable and postpone the development of new business models. With an increased bargaining power of these vendors, the success of initiatives such as *SMART on FHIR* may emerge from the tension between providers' needs and vendors' desire to keep control over their products [19].

Some researchers have suggested that the use of similar systems across the country will create opportunities for human factors researchers by facilitating comparison of similar functionality [5]; however, such opportunities may not reach fruition because of local configurations that allow the same product to be implemented in completely different ways across clients [32]. Overcoming the vendor oligopoly will demand development of informatics solutions proved to be more effective than current systems' functionality, which leads us to the next UC: innovation vacuum.

Unintended Consequence 3: Innovation Vacuum

As EHR adoption has primarily been achieved through financial incentives, the cycle of technological innovation typical of other industries has not been observed in the US HIT sector. As a result, commercial EHRs were adopted before fixing widely known problems such as poor usability [33], which has been associated to patient harm [34,35], and suboptimal clinical decision support (CDS) systems [36] such as excessive, overzealous alerts frequently ignored by providers [37]. In addition, a recent evaluation of EHR certification criteria concluded that the certification process is not designed to prevent patient harm [38]. Specifically, the report found that the usability testing required does not include a representative sample, does not include real clinical scenarios, and does not simulate changes added through system configuration by local clients.

The accelerated adoption also affected benchmarking organizations such as Intermountain Healthcare, Partners Healthcare, and the Veterans Health Administration that have traditionally promoted most HIT innovations [39]. These organizations decided to replace their systems with commercial EHRs, putting an end to the homegrown systems' era. As a result, some of these organizations decided to dissolve their informatics departments [40,41], decreasing their investment in informatics innovation.

XSL•FO
RenderX

With widespread adoption of suboptimal and poorly tested systems, along with traditional innovators stepping aside, fixing the EHR now is a bit like fixing an airplane midflight, and without a pilot.

The Path Forward

At least 2 panels at recent American Medical Informatics Association annual symposia have presented informatics innovations in the post-MU era with clients of 1 large HIT vendor, and most innovations included SMART on FHIR apps [42,43]. Panelists have pointed out that as commercial EHRs can properly handle capabilities such as billing, data storage, and privacy regulations, informatics innovators tend to be freer to innovate in the post-MU era. However, as previously mentioned, most HIT vendors are not yet fully open to seamless interface with external apps. In addition, FHIR is a standard under development, and a substitute for the traditional innovators is yet to be found. To aggravate the problem, most contracts signed between providers and HIT vendors include clauses that hamper transparency by preventing providers from sharing usability and safety issues that could otherwise advance EHR design [44].

There was a natural reason for having most HIT innovations coming from benchmarking organizations: neither HIT vendors nor academic departments have seamless access to clinicians at the point of care, where informatics applications are put to the test. In naturalistic settings, iterations between clinicians and informaticists facilitate an understanding of users' needs to inform EHR development. Academic informatics departments could serve as a natural replacement for the traditional innovators by promoting cutting-edge research toward fixing the EHR, coupled with more robust HIT evaluations. However, this replacement will demand a closer relationship between academic departments and their medical centers. In US universities, these departments tend to function as independent organizations, which hampers researchers' access to HIT resources and clinicians at the point of care. Work in such a direction has started [45-47] and serves as example of the path needed to design new business models, fostering innovation and transparency, and fixing the EHR.

Unintended Consequence 4: Physician Burnout

The accelerated adoption of commercial EHRs coincided (and likely was programmed to coincide) with the implementation of the Affordable Care Act (ACA). The slow, but steady, implementation of pay-for-performance payment models has given rise to the EHR-based quality measurement [48]. The push for reporting clinical performance generates an increased demand for capturing accurate, structured data [5], and the use of suboptimal EHRs in these tasks has contributed to the so called EHR-associated physician burnout [49]. The use of clinical documentation for nonclinical purposes is increasing and is source of frustration among physicians [50,51]. This is reinforced by the fact that electronic clinical notes generated in the United States are significantly longer than similar documentation in other developed countries [52]. Recent studies have found that in the post-MU and ACA era, for every hour of patient contact time, physicians may spend up to 2 hours on electronic documentation [53,54]. The documentation burden has been so intense that in some cases, physicians intentionally close slots in their agenda to complete electronic documentation of previous patients [17].

The Path Forward

In addition to simplifying billing requirements [6] and developing informatics solutions to extract quality indicators from clinical documentation [5], a fundamental redesign of the EHR to improve data entry and retrieval is needed. The structured and static format of current EHR interfaces force physicians to record clinical data through predefined and strict functionality dependent on the current *desktop kit* (pointer + keyboard + monitor with a cluttered EHR interface). For physicians to keep the richer narrative of their clinical assessments while decreasing the documentation burden, EHRs must demand less typing and clicking [55]. New technologies such as conversational speech recognition (CSR) have recently achieved human parity with regards to transcription error rate [56] and have tremendous potential for substantially decreasing typing and clicking. However, CSR solutions may be compromised by the fact that clinicians may make conscious decisions about what information to communicate to patients and to document in the EHR [57]. Therefore, there are opportunities for research exploring what information clinicians document (or not) in the EHR and what information they do not communicate verbally to the patient but document in their clinical notes [58]; such findings will inform development of CSR and other data-entry solutions capable of handling such situations. Regarding data retrieval, EHR content retrieved by physicians is influenced by their tasks or information goals [59,60]; however, such stimuli are not captured by current EHRs. Future research should investigate how EHRs can support data retrieval with intelligent stimulus- or goal-oriented functionality that allows a holistic view of the patient and flexible navigation across the record [58] to hopefully decrease the documentation burden and its contribution to the next UC: data obfuscation.

Unintended Consequence 5: Data Obfuscation

Physicians frequently create their clinical notes by using the patient's previous note, a practice known as *copy-and-paste*. [61] As a result, they often produce (and later deal with) uninformative, bloated notes that often contain redundant information and errors [62,63]. In addition, these notes do not provide the data in a way that increases clinicians' situational awareness (ie, the perception and comprehension of relevant information necessary to take action) [64], and in some cases may never be read [65]. The problem is aggravated by overwhelming CDS alerts and reminders; many clinicians complain that such alerts make them vulnerable to information overload, which might lead them to miss important information [66]. The obfuscation of relevant data resulting from bloated records has been reported [67,68], associated with potential

XSL•FO
RenderX

safety hazards [69] and with delayed or incorrect decisions at the point of care [70].

The Path Forward

Some proposed solutions to highlighting relevant data include tailoring physicians' use of EHRs to document *what they are thinking* about the patient's situation [64], transferring some data entry to patients [6], or new policies to facilitate health information exchange (HIE) [5,6,71]. Such proposals are unlikely to succeed in isolation as they require clinicians to enter or import even more information into already bloated records. In addition, the effectiveness of HIE seems to be understudied [72]; although some studies report HIE-associated improvements [73], others report the opposite [74].

Concise documentation that highlights relevant data will come from smarter EHRs that actively participate in patient care [75]; however, to be smarter, EHRs must be able to capture and process more information about the patient's context and clinicians' reasoning. Previous studies suggest that clinicians seem to always know something that is only partially represented in or is missing entirely from the EHR [37,76]. For example, EHRs are incapable of understanding why clinicians order what they order, or how current symptoms are related to previous problems. Although most EHRs allow medical records to be structured on a problem-oriented basis, such structure does not capture the reasoning behind the relationship between problems and other clinical concepts. For example, a medication can be linked to a problem, indicating that it was ordered to treat a particular problem, but the reasoning (why) behind the choice for this particular medication is not captured by the EHR. If such data were captured, several opportunities for informatics research would emerge to apply (and improve) computational methods (eg, machine learning, natural language processing, and text generation methods) to empower the EHR to use patient's care context data. Context-rich data could be used to facilitate note creation, to create automatic notes ready for review, and to increase the accuracy of CDS, potentially mitigating the already infamous alert fatigue [37]. However, 2 major challenges remain: (1) A formal representation of the semantic relationships between clinical concepts (eg, symptoms, findings, problems, diagnoses, and treatments) does not exist and (2) Effective methods for capturing and representing clinicians' reasoning need to be developed [58]. EHR vendors have avoided this path to avert coliability for medical errors when eventual system failures lead to misleading recommendations [77,78]. What vendors have avoided translates into several opportunities for informatics researchers. The development of a formal representation of clinicians reasoning seems to be a promising alternative to empower EHRs to represent patients' situation [79]. However, the application of such a representation into actual patient data will demand new, more effective data-entry approaches [58], improvements to data visualization [80], and computational methods [55].

On balance, despite the unexpected effects and challenges of nationwide EHR adoption, several opportunities for developing more effective EHRs and evaluation methods are likely to emerge from the forces promoting progress. The UCs here discussed do not intend to be exhaustive; other consequences may be revealed as new, more robust HIT evaluations are reported. We hypothesize that overcoming these UCs will likely require a path reverse to the one that produced them. By creating smarter clinical information systems with more intuitive navigation and data entry functionality, clinicians could save time searching, synthesizing, and documenting data in the EHR, which would contribute to alleviate data obfuscation and mitigate burnout. Such systems will likely come from external applications developed through cutting-edge research conducted in academic medical centers that tend to be a natural replacement for earlier informatics innovators. These applications, if successfully implemented and evaluated, may back providers up on their demands to have most large EHR vendors opening their platforms, which would facilitate the development of new business models and decrease market oligopoly. Finally, by accumulating evidence of the effectiveness of these applications, in isolation and in conjunction with commercial EHRs, a better understanding of the true positive effects of HIT can be obtained by future systematic reviews and meta-analyses.

The multiple efforts proposed here will demand collaboration between diverse players such as health care providers, administrators, HIT vendors, policy makers, informatics researchers, funding agencies, and outside developers toward a single goal: to realize the full potential of a digitized health system.

Acknowledgments

This work was supported by research funds from the Informatics Institute of the University of Alabama at Birmingham.

Conflicts of Interest

None declared.

References

1. Garfield E. The unintended and unanticipated consequences of Robert K Merton. Soc Stud Sci 2016 Jun 29;34(6):845-853. [doi: 10.1177/0306312704042087]
2. Institute of Medicine. The Computer-based Patient Record: An Essential Technology for Health Care. Revised Edition. Dick RS, Steen EB, Detmer DE. editors. Washington, DC: National Academy Press; 1997.
3. United States Department of Health and Human Services. HITECH Act § 13400(5) URL: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf [accessed 2019-05-10] [WebCite Cache ID 78H4DGfp7]

XSL•FO
RenderX

4. The Office of the National Coordinator for Health Information Technology. Health IT Dashboard URL: https://dashboard.healthit.gov/quickstats/quickstats.php [accessed 2019-05-10] [WebCite Cache ID 78H1fzXpZ]
5. Sittig DF, Wright A, Ash J, Singh H. New unintended adverse consequences of electronic health records. Yearb Med Inform 2016 Nov 10(1):7-12. [doi: 10.15265/IY-2016-023] [Medline: 27830226]
6. Payne TH, Corley S, Cullen TA, Gandhi TK, Harrington L, Kuperman GJ, et al. Report of the AMIA EHR-2020 Task Force on the status and future direction of EHRs. J Am Med Inform Assoc 2015 Sep;22(5):1102-1110. [doi: 10.1093/jamia/ocv066] [Medline: 26024883]
7. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471 [FREE Full text] [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
8. Jones SS, Rudin RS, Perry T, Shekelle PG. Health information technology: an updated systematic review with a focus on meaningful use. Ann Intern Med 2014 Jan 7;160(1):48-54. [doi: 10.7326/M13-1531] [Medline: 24573664]
9. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood) 2005;24(5):1103-1117 [FREE Full text] [doi: 10.1377/hlthaff.24.5.1103] [Medline: 16162551]
10. Papanicolas I, Woskie LR, Jha AK. Health care spending in the United States and other high-income countries. J Am Med Assoc 2018 Mar 13;319(10):1024-1039. [doi: 10.1001/jama.2018.1150] [Medline: 29536101]
11. Colicchio TK, Fiol GD, Stoddard GJ, Narus SP. Evaluation of a systematic methodology to detect in near real-time performance changes during electronic health record system implementations: a longitudinal study. AMIA Annu Symp Proc 2017;2017:595-604 [FREE Full text] [Medline: 29854124]
12. Colicchio TK, Del Fiol G, Scammon DL, Facelli JC, Bowes WA, Narus SP. Comprehensive methodology to monitor longitudinal change patterns during EHR implementations: a case study at a large health care delivery network. J Biomed Inform 2018 Jul;83:40-53. [doi: 10.1016/j.jbi.2018.05.018] [Medline: 29857137]
13. Murphy DR, Meyer AN, Russo E, Sittig DF, Wei L, Singh H. The burden of inbox notifications in commercial electronic health records. JAMA Intern Med 2016 Apr;176(4):559-560 [FREE Full text] [doi: 10.1001/jamainternmed.2016.0209] [Medline: 26974737]
14. Bryant AD, Fletcher GS, Payne TH. Drug interaction alert override rates in the Meaningful Use era: no evidence of progress. Appl Clin Inform 2014;5(3):802-813 [FREE Full text] [doi: 10.4338/ACI-2013-12-RA-0103] [Medline: 25298818]
15. Kruse CS, Beane A. Health information technology continues to show positive effect on medical outcomes: systematic review. J Med Internet Res 2018 Feb 05;20(2):e41 [FREE Full text] [doi: 10.2196/jmir.8793] [Medline: 29402759]
16. Brynjolfsson E, Hitt LM. Beyond the productivity paradox. Commun ACM 1998 Aug;41(8):49-55. [doi: 10.1145/280324.280332]
17. Colicchio TK, Borbolla D, Colicchio VD, Del Fiol G, Scammon DL, Facelli JC, et al. Looking Behind the Curtain: Identifying Factor Affecting Care Outcomes During a Large Commercial EHR Implementation. eGEMs 2019 May 06;7((1): 21):1-12 [FREE Full text]
18. Jones SS, Heaton PS, Rudin RS, Schneider EC. Unraveling the IT productivity paradox--lessons for health care. N Engl J Med 2012 Jun 14;366(24):2243-2245. [doi: 10.1056/NEJMp1204980] [Medline: 22693996]
19. Wachter RM, Howell MD. Resolving the productivity paradox of health information technology: a time for optimism. J Am Med Assoc 2018 Jul 03;320(1):25-26. [doi: 10.1001/jama.2018.5605] [Medline: 29813165]
20. Colicchio TK, Facelli JC, Del Fiol G, Scammon DL, Bowes WA, Narus SP. Health information technology adoption: understanding research protocols and outcome measurements for IT interventions in health care. J Biomed Inform 2016 Dec;63:33-44 [FREE Full text] [doi: 10.1016/j.jbi.2016.07.018] [Medline: 27450990]
21. Ioannidis JP. Why most published research findings are false. PLoS Med 2005 Aug;2(8):e124 [FREE Full text] [doi: 10.1371/journal.pmed.0020124] [Medline: 16060722]
22. Koppel R. Is healthcare information technology based on evidence? Yearb Med Inform 2013;8:7-12. [Medline: 23974542]
23. Rudin RS, Jones SS, Shekelle P, Hillestad RJ, Keeler EB. The value of health information technology: filling the knowledge gap. Am J Manag Care 2014 Nov;20(11 Spec No. 17):eSP1-eSP8 [FREE Full text] [Medline: 25811814]
24. Harrison MI, Koppel R, Bar-Lev S. Unintended consequences of information technologies in health care--an interactive sociotechnical analysis. J Am Med Inform Assoc 2007;14(5):542-549 [FREE Full text] [doi: 10.1197/jamia.M2384] [Medline: 17600093]
25. Lagarde M. How to do (or not to do) ... Assessing the impact of a policy change with routine longitudinal data. Health Policy Plan 2012 Jan;27(1):76-83. [doi: 10.1093/heapol/czr004] [Medline: 21278077]
26. Hacker K, Penfold R, Zhang F, Soumerai SB. Impact of electronic health record transition on behavioral health screening in a large pediatric practice. Psychiatr Serv 2012 Mar;63(3):256-261 [FREE Full text] [doi: 10.1176/appi.ps.201100207] [Medline: 22267253]
27. Joseph S, Sow M, Furukawa MF, Posnack S, Chaffee MA. HITECH spurs EHR vendor competition and innovation, resulting in increased adoption. Am J Manag Care 2014 Sep;20(9):734-740 [FREE Full text] [Medline: 25365748]

XSL•FO
RenderX

28. HIMSS Analytics. Which EHR vendor boosts the most hospital clients URL: https://www.beckershospitalreview.com/healthcare-information-technology/himss-analytics-which-ehr-vendor-boasts-the-most-hospital-clients.html [accessed 2019-05-10] [WebCite Cache ID 78H1uhMHv]
29. KLAS Research. Hospital EMR Market Share URL: https://klasresearch.com/report/us-hospital-emr-market-share-2018/1279 [accessed 2019-05-10] [WebCite Cache ID 78H2NA96Y]
30. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Sep;23(5):899-908 [FREE Full text] [doi: 10.1093/jamia/ocv189] [Medline: 26911829]
31. KLAS Research. EMR Global Market URL: https://klasresearch.com/report/global-emr-market-share-2017/1192 [accessed 2019-05-10] [WebCite Cache ID 78H2R8Sxz]
32. Koppel R, Lehmann CU. Implications of an emerging EHR monoculture for hospitals and healthcare systems. J Am Med Inform Assoc 2015 Mar;22(2):465-471. [doi: 10.1136/amiajnl-2014-003023] [Medline: 25342181]
33. Middleton B, Bloomrosen M, Dente MA, Hashmat B, Koppel R, Overhage JM, et al. Enhancing patient safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA. J Am Med Inform Assoc 2013 Jun;20(e1):e2-e8 [FREE Full text] [doi: 10.1136/amiajnl-2012-001458] [Medline: 23355463]
34. Howe JL, Adams KT, Hettinger AZ, Ratwani RM. Electronic health record usability issues and potential contribution to patient harm. J Am Med Assoc 2018 Dec 27;319(12):1276-1278 [FREE Full text] [doi: 10.1001/jama.2018.1171] [Medline: 29584833]
35. Ratwani RM, Savage E, Will A, Fong A, Karavite D, Muthu N, et al. Identifying electronic health record usability and safety challenges in pediatric settings. Health Aff (Millwood) 2018 Dec;37(11):1752-1759. [doi: 10.1377/hlthaff.2018.0699] [Medline: 30395517]
36. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. AMIA Annu Symp Proc 2007:26-30 [FREE Full text] [Medline: 18693791]
37. Baysari MT, Tariq A, Day RO, Westbrook JI. Alert override as a habitual behavior-a new perspective on a persistent problem. J Am Med Inform Assoc 2017 Dec 01;24(2):409-412. [doi: 10.1093/jamia/ocw072] [Medline: 27274015]
38. The Pew Trusts. Ways to Improve Electronic Health Record Safety URL: https://www.pewtrusts.org/en/research-and-analysis/reports/2018/08/28/ways-to-improve-electronic-health-record-safety [accessed 2019-05-10] [WebCite Cache ID 78H2Uya7M]
39. Evans RS. Electronic health records: then, now, and in the future. Yearb Med Inform 2016 May 20;Suppl 1:S48-S61. [doi: 10.15265/IYS-2016-s006] [Medline: 27199197]
40. Bowes WA. Impacts of EHR certification and meaningful use implementation on an integrated delivery network. AMIA Annu Symp Proc 2014;2014:325-332 [FREE Full text] [Medline: 25954335]
41. DeseretNews. Intermountain Healthcare to Outsource 98 Information Technology Employees URL: https://www.deseretnews.com/article/900014140/98-intermountain-healthcare-information-technology-employees-being-outsourced.html [accessed 2019-05-10] [WebCite Cache ID 78H4Ngvot]
42. AMIA Annual Symposium 2015. URL: https://www.amia.org/amia2015/panels [accessed 2019-05-10] [WebCite Cache ID 78H2dsODP]
43. AMIA Annual Symposium 2018. Agenda URL: https://symposium2018.zerista.com/event/member/507934 [accessed 2019-05-10] [WebCite Cache ID 78H2jJk8N]
44. Ratwani R, Hodgkins M, Bates D. Improving electronic health record usability and safety requires transparency. J Am Med Assoc 2018 Dec 25;320(24):4. [Medline: 30489619]
45. Crump J, Del Fiol G, Williams M, Freimuth R. Prototype of a Standards-Based EHR and Genetic Test Reporting Tool Coupled with HL7-Compliant Infobuttons. AMIA Jt Summits Transl Sci Proc 2018;2017:330-339 [FREE Full text] [Medline: 29888091]
46. Bergquist T, Buie R, Li K, Brandt P. Heart on FHIR: integrating patient generated data into clinical care to reduce 30 day heart failure readmissions (Extended Abstract). AMIA Annu Symp Proc 2017;2017:2269-2273 [FREE Full text] [Medline: 29854266]
47. Sinha S, Jensen M, Mullin S, Elkin P. Safe Opioid Prescription: a SMART on FHIR approach to clinical decision support. Online J Public Health Inform 2017 Sep 08;9(2):A. [Medline: 29026458]
48. Cusack CM, Hripcsak G, Bloomrosen M, Rosenbloom ST, Weaver CA, Wright A, et al. The future state of clinical data capture and documentation: a report from AMIA's 2011 Policy Meeting. J Am Med Inform Assoc 2013 Jan 1;20(1):134-140 [FREE Full text] [doi: 10.1136/amiajnl-2012-001093] [Medline: 22962195]
49. Shanafelt TD, Dyrbye LN, West CP. Addressing physician burnout: the way forward. J Am Med Assoc 2017 Dec 07;317(9):901-902. [doi: 10.1001/jama.2017.0076] [Medline: 28196201]
50. Ho Y, Gadd C, Kohorst K, Rosenbloom S. A qualitative analysis evaluating the purposes and practices of clinical documentation. Appl Clin Inform 2014;5(1):153-168 [FREE Full text] [doi: 10.4338/ACI-2013-10-RA-0081] [Medline: 24734130]
51. Koopman R, Steege L, Moore J, Clarke M, Canfield S, Kim M, et al. Physician information needs and Electronic Health Records (EHRs): time to reengineer the clinic note. J Am Board Fam Med 2015;28(3):316-323 [FREE Full text] [doi: 10.3122/jabfm.2015.03.140244] [Medline: 25957364]

XSL•FO
RenderX

52. Downing N, Bates D, Longhurst C. Physician burnout in the electronic health record era. Ann Intern Med 2019 Feb 05;170(3):216-217. [doi: 10.7326/L18-0604] [Medline: 30716744]
53. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of Physician Time in Ambulatory Practice: A Time and Motion Study in 4 Specialties. Ann Intern Med 2016 Dec 06;165(11):753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]
54. Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan W, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med 2017 Sep;15(5):419-426 [FREE Full text] [doi: 10.1370/afm.2121] [Medline: 28893811]
55. Wachter R, Goldsmith J. Harvard Business Review. 2018. To combat physician burnoutimprove care, fix the electronic health record URL: https://hbr.org/2018/03/to-combat-physician-burnout-and-improve-care-fix-the-electronic-health-record [accessed 2019-05-14] [WebCite Cache ID 78MyJabp4]
56. Xiong W, Droppo J, Huang X, Seide F, Seltzer M, Stolcke A. arXivcs Internet. 2016. Achieving Human Parity in Conversational Speech Recognition URL: http://arxiv.org/abs/1610.05256 [accessed 2018-04-25]
57. Tuepker A, Zickmund S, Nicolajski C, Hahm B, Butler J, Weir C. Providers? Note-Writing Practices for Post-traumatic Stress Disorder at Five United States Veterans Affairs Facilities. J Behav Health Serv Res 2016 Jul 01;43(3):42. [Medline: 26228929]
58. Colicchio T, Cimino JJ. Clinicians' reasoning as reflected in electronic clinical note-entry and reading/retrieval: a systematic review and qualitative synthesis. J Am Med Inform Assoc 2019 Feb 01;26(2):172-184. [doi: 10.1093/jamia/ocy155] [Medline: 30576561]
59. Rizvi RF, Harder KA, Hultman GM, Adam TJ, Kim M, Pakhomov SV, et al. A comparative observational study of inpatient clinical note-entry and reading/retrieval styles adopted by physicians. Int J Med Inform 2016 Jun;90:1-11. [doi: 10.1016/j.ijmedinf.2016.02.011] [Medline: 27103191]
60. Wright M, Dunbar S, Macpherson B, Moretti E, Del Fiol G, Bolte J, et al. Toward designing information display to support critical care. A qualitative contextual evaluation and visioning effort. Appl Clin Inform 2016 Dec 05;7(4):912-929 [FREE Full text] [doi: 10.4338/ACI-2016-03-RA-0033] [Medline: 27704138]
61. Siegler EL, Adelman R. Copy and paste: a remediable hazard of electronic health records. Am J Med 2009 Jun;122(6):495-496. [doi: 10.1016/j.amjmed.2009.02.010] [Medline: 19486708]
62. Wrenn J, Stein D, Bakken S, Stetson P. Quantifying clinical narrative redundancy in an electronic health record. J Am Med Inform Assoc 2010;17(1):49-53 [FREE Full text] [doi: 10.1197/jamia.M3390] [Medline: 20064801]
63. Hammond KW, Helbig ST, Benson CC, Brathwaite-Sketoe BM. Are electronic medical records trustworthy? Observations on copying, pasting and duplication. AMIA Annu Symp Proc 2003:269-273 [FREE Full text] [Medline: 14728176]
64. Lenert L. Toward medical documentation that enhances situational awareness learning. AMIA Annu Symp Proc 2016;2016:763-771 [FREE Full text] [Medline: 28269872]
65. Hripcsak G, Vawdrey DK, Fred MR, Bostwick SB. Use of electronic clinical documentation: time spent and team interactions. J Am Med Inform Assoc 2011;18(2):112-117 [FREE Full text] [doi: 10.1136/jamia.2010.008441] [Medline: 21292706]
66. Singh H, Spitzmueller C, Petersen NJ, Sawhney MK, Sittig DF. Information overload and missed test results in electronic health record-based settings. JAMA Intern Med 2013 Apr 22;173(8):702-704 [FREE Full text] [doi: 10.1001/2013.jamainternmed.61] [Medline: 23460235]
67. Tsou A, Lehmann C, Michel J, Solomon R, Possanza L, Gandhi T. Safe practices for copy and paste in the EHR. Systematic review, recommendations, and novel model for health IT collaboration. Appl Clin Inform 2017 Dec 11;8(1):12-34 [FREE Full text] [doi: 10.4338/ACI-2016-09-R-0150] [Medline: 28074211]
68. Partnership for Health IT Patient Safety. 2016 Feb. Health IT Safe Practices: Toolkit for the Safe Use of Copy and Paste URL: https://www.ecri.org/Resources/HIT/CP_Toolkit/Toolkit_CopyPaste_final.pdf [accessed 2019-05-10] [WebCite Cache ID 78H3G7AoQ]
69. Bowman S. Impact of electronic health record systems on information integrity: quality and safety implications. Perspect Health Inf Manag 2013;10:1c [FREE Full text] [Medline: 24159271]
70. Singh H, Giardina T, Meyer A, Forjuoh S, Reis M, Thomas E. Types and origins of diagnostic errors in primary care settings. JAMA Intern Med 2013 Mar 25;173(6):418-425 [FREE Full text] [doi: 10.1001/jamainternmed.2013.2777] [Medline: 23440149]
71. Halamka JD, Tripathi M. The HITECH era in retrospect. N Engl J Med 2017 Sep 07;377(10):907-909. [doi: 10.1056/NEJMp1709851] [Medline: 28877012]
72. Rahurkar S, Vest JR, Menachemi N. Despite the spread of health information exchange, there is little evidence of its impact on cost, use, and quality of care. Health Aff (Millwood) 2015 Mar;34(3):477-483. [doi: 10.1377/hlthaff.2014.0729] [Medline: 25732499]
73. Frisse ME, Johnson KB, Nian H, Davison CL, Gadd CS, Unertl KM, et al. The financial impact of health information exchange on emergency department care. J Am Med Inform Assoc 2012;19(3):328-333 [FREE Full text] [doi: 10.1136/amiajnl-2011-000394] [Medline: 22058169]

XSL•FO
RenderX

74. Lang E, Afilalo M, Vandal AC, Boivin J, Xue X, Colacone A, et al. Impact of an electronic link between the emergency department and family physicians: a randomized controlled trial. Can Med Assoc J 2006 Jan 31;174(3):313-318 [FREE Full text] [doi: 10.1503/cmaj.050698] [Medline: 16399880]
75. Cimino JJ. Improving the electronic health record--are clinicians getting what they wished for? J Am Med Assoc 2013 Mar 13;309(10):991-992 [FREE Full text] [doi: 10.1001/jama.2013.890] [Medline: 23483171]
76. Her Q, Amato M, Seger D, Beeler P, Slight S, Dalleur O, et al. The frequency of inappropriate nonformulary medication alert overrides in the inpatient setting. J Am Med Inform Assoc 2016 Sep;23(5):924-933. [Medline: 27002076]
77. Koppel R, Kreda D. Health care information technology vendors' "hold harmless" clause: implications for patients and clinicians. J Am Med Assoc 2009 Mar 25;301(12):1276-1278. [doi: 10.1001/jama.2009.398] [Medline: 19318655]
78. Wachter R. The Digital Doctor: hope, hype, and harm at the dawn of medicine?. The Digital Doctor: McGraw-Hill, New York; 2015:1260019608-1260019690.
79. Cimino J, Li Z, Weng C. An exploration of the terminology of clinical cognition and reasoning. AMIA Annu Symp Proc 2018;2018:321-329 [FREE Full text] [Medline: 30815071]
80. Nolan ME, Siwani R, Helmi H, Pickering BW, Moreno-Franco P, Herasevich V. Health IT usability focus section: data use and navigation patterns among medical ICU clinicians during electronic chart review. Appl Clin Inform 2017 Dec;8(4):1117-1126. [doi: 10.4338/ACI-2017-06-RA-0110] [Medline: 29241249]

Abbreviations

ACA: Affordable Care Act
ARRA: America Recovery and Reinvestment Act
CDS: clinical decision support
CSR: conversational speech recognition
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
HIE: health information exchange
HIT: health information technology
HITECH: Health Information Technology for Economic and Clinical Health
MU: Meaningful Use
SMART: Substitutable Medical Applications & Reusable Technologies
UCs: unintended consequences

Edited by G Eysenbach; submitted 07.01.19; peer-reviewed by A Martínez-García, N Delvaux, B Vaes; comments to author 28.03.19; revised version received 09.04.19; accepted 26.04.19; published 03.06.19.

Please cite as:
Colicchio TK, Cimino JJ, Del Fiol G
Unintended Consequences of Nationwide Electronic Health Record Adoption: Challenges and Opportunities in the Post-Meaningful Use Era
J Med Internet Res 2019;21(6):e13313
URL: https://www.jmir.org/2019/6/e13313/
doi:10.2196/13313
PMID:31162125

©Tiago K Colicchio, James J Cimino, Guilherme Del Fiol. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Toward Comprehensive Patient-Centric Care by Integrating Digital Health Technology With Direct Clinical Contact in Australia

Penelope Schofield[1,2,3,4], BSc (Hons), PhD; Tim Shaw[5], BSc (Hons), PhD; Michaela Pascoe[2,6], BSc Psych (Hons), PhD

[1]Department of Psychology, Swinburne University, Melbourne, Australia

[2]Department of Cancer Experiences Research, Peter MacCallum Cancer Centre, Melbourne, Australia

[3]Sir Peter MacCallum Department of Oncology, The University of Melbourne, Melbourne, Australia

[4]Iverson Health Innovation Research Institute, Swinburne University, Melbourne, Australia

[5]Charles Perkins Centre, University of Sydney, Sydney, Australia

[6]The Institute for Health and Sport, Victoria University, Melbourne, Australia

Corresponding Author:
Penelope Schofield, BSc (Hons), PhD
Sir Peter MacCallum Department of Oncology
The University of Melbourne
Parkville VIC
Melbourne, 3010
Australia
Phone: 61 03 9214 4886 ext 4886
Fax: 61 03 9214 4886
Email: pschofield@swin.edu.au

Abstract

Background: There is an escalating crisis in health care, locally and internationally. The current health care model is unable to meet the increasing health care demands.

Objective: The aim of this study was to reconceptualize the provision of health care to produce better outcomes at no greater cost, by placing individuals in the position of authority to direct their own care, in a personalized, integrated health care system.

Methods: In this study, we used the Australian health care system as a model. We reviewed the current landscape of digital health in Australia and discussed how electronic medical records (EMRs) can be further developed into a personalized, integrated health care system.

Results: Some components of an EMR and digital health system are already being used in Australia, but the systems are not linked. A personalized, integrated health care model that is responsive to consumer needs requires not just a passive repository of medical information; it would require a team approach, including the government, health care funders, industries, consumers and advocacy groups, health care professionals, community groups, and universities.

Conclusions: Implementation of a personalized, integrated health care system can result in reduced pressure on the current health care system, and it can result in the delivery of best-practice health care, regardless of location. Importantly, a personalized, integrated health care system could serve as an education platform, “upskilling” not only clinicians but also, more importantly, patients and carers by providing them with accurate information about their condition, treatment options, medications, and management strategies. By proposing personalized, integrated health care, we offer an intelligent model of health care that is ubiquitous, efficient, and continuously improving.

(J Med Internet Res 2019;21(6):e12382) doi:10.2196/12382

KEYWORDS

health care; health; eHealth

XSL•FO
RenderX

Introduction

Background

Ongoing improvements in health care, successful management of chronic conditions, and falling birth rates in developed nations are globally leading to an aging population [1]. In 1950, 1 person in every 12 people was aged above 60 years [2], by 2015, this had risen to 1 person in every 8 people worldwide [3]. The proportion of older persons is projected to continue to increase, with 1 in 5 people estimated to be aged above 60 by 2050 [3]. The associated prevalence of chronic diseases is immense [1]. By 2020, noncommunicable diseases are expected to contribute to 57% of the global burden of disease and 75% of deaths [4]. Indeed, approximately 40 million people died because of a noncommunicable disease in 2015, namely cardiovascular disease, cancer, chronic respiratory disease, and diabetes, accounting for 70% of deaths worldwide [5]. The societal burden of managing this growing health crisis is substantial and unable to be supported by current models of health care. Like many other countries, Australia is turning to digital technologies to address the rapidly growing gap between service demand and capacity. Achieving a sustainable, agile, and effective health system that keeps pace with demands requires a fundamental disruption of health care delivery. The health system will need to shift focus away from health care providers in centralized locations treating illness; instead, the health system will need to shift focus toward supporting empowered consumers in the community in managing their own health and well-being.

In Australia, the gross domestic product expenditure on health care has increased from 6.5% in 1989-90 to 10.3% in 2015-16 [6,7]. Annually, in Australia, there are an estimated 230,000 medication-related hospital admissions, with a cost to taxpayers of Aus $1.2 billion [8]. The Australian Medicare system is a publicly funded, national universal health care system. In the 2016-17 financial year, Australian Medicare expenditure on medications was Aus $12 million, an increase of 11.3% from the previous financial year [9]. This economic burden is exacerbated by inefficiencies and waste in health care provision. Bentley et al (2008) [10] outlined 4 key inefficiencies: (1) duplication of services, such as repeated blood tests, as clinicians have no access to results of tests ordered by other clinicians; (2) inefficient processes, such as unnecessary transport of patients to seek medical management not available locally; (3) overly expensive inputs, such as physicians providing services that nurses are equally competent to provide; (4) medical errors, such as adverse medication events resulting in avoidable hospital admissions.

The downstream consequence of increasing demand on the health system, errors, inefficiency, and waste is clinician burnout. Like many countries, Australia has a workforce shortage across a number of health professions [11]; therefore, clinicians are struggling to meet the health care requirements of patients [12]. Exacerbating this situation, consumers have increased expectations of clinicians to assist patients in becoming more informed and involved in the treatment decision making and health management and provide prompt, appropriate, and individualized support [12]. Unsurprisingly, high rates of burnout and psychological distress are experienced by medical doctors: indeed, 48% of younger doctors report burnout in the domain of emotional exhaustion, 18% report low professional efficacy, and 46% report high cynicism, according to a survey of over 12,000 Australian medical doctors [13]. Clinician burnout not only greatly reduces quality of life for clinicians but also reduces quality of care to patients and contributes to workforce shortages [14].

The current political climate in health care is characterized by fiscal restraints combined with an aging population [3], increases in the incidence of chronic illnesses [4], and escalating health care costs [6]. These issues are aggravated by a limited and burnt-out workforce trying to meet growing needs [11,14] and significant inefficacies and waste [10]. The existing system of health care is an insatiable beast, which we need to change. Internationally, efforts are presently focused on translating paper-based records into electronic records. In many respects, this represents a digitization of eighteenth century health care and misses a crucial opportunity for technology to transform care and support new disruptive models of care [15]. Currently, routine digital technology use in health is fragmented and mostly “passive” in assisting with information storage or analysis. Although technology heavy, health care has undertaken a digital transformation in the way that businesses, particularly financial services, have. To ensure sustainable effective health care into the future, it is crucial that we implement a digitally supported, fully integrated, and secure health care system that is proactive in the following: tailoring precision, holistic health solutions for the individual; improving equity and access, particularly among the disadvantaged; disseminating health innovations rapidly and upskilling health providers in these innovations; empowering consumers to manage their own health and well-being [16].

Objectives

The aim of this paper was to reconceptualize the provision of health care to produce better outcomes at no greater cost, by placing individuals in the position of authority to direct their own care. The fundamental premise is that individuals are the experts in their illness experience and personal values in relation to the optimal care they require. In this paper, we use the Australian health care system as a model. We first introduce the concept of a “Bespoke Health Care System” (BHS) and discuss how digital health is currently being implemented in the Australian health care system, including the use and potential of electronic medical records (EMRs). We then propose that EMRs could be integrated into a patient-centered management platform rather than just function as a linked passive information depository. Finally, the necessary conditions for the successful adoption of bespoke health care are discussed, including codesign to develop a platform for the user to ensure positive user experience and trust in the data security and the highlight the importance of utilizing frameworks to identify the barriers most relevant to the implementation BHS.

Bespoke Health Care

Modern education has adopted an effective new approach of “flipping the classroom,” whereby application of knowledge is done in the classroom whereas didactic learning is self-directed

XSL•FO
RenderX

and occurs before class [17,18]. A number of health care practitioners are applying "flipping" to clinical health care, that is, providing patients with necessary knowledge before consultations and then using consultation time to problem solve and make joint decisions, with the aim of improving health outcomes and experience of care and reducing costs [19].

We propose a new health care system that builds upon the flipped health care model to place consumers in the driver's seat of their health care, termed the BHS. A central component of this model is increased patient involvement in health care decisions and self-management assisted by the use of technology. In the proposed system, individuals will have an EMR, which contains all of a person's relevant personal and medical information across providers. A total of 1 in 5 medical errors are because of incomplete patient medication information [20]. This record could be linked to primary care clinics, specialists' rooms, public and private hospitals, pathology services, pharmacies, insurance companies, and other relevant health agencies. It would allow new and relevant medical, diagnostic, and management information to be uploaded and viewed by the individual's network of health care providers in real time. Where appropriate, individuals would have access to their medical records via the EMR by using their computers and smartphones.

However, the BHS expands well beyond the use of traditional EMRs. Importantly, an EMR need not be a passive depository of data. It could be used to remotely track symptoms, provide routine medication reminders, prompt patients to have a routine screening test, such as a breast mammogram, or allow patients to book medical appointments and remind them to attend. Currently, patient records are still often paper based. Even when they are electronic, they are often not linked, disorganized, and unable to be easily accessed by patients or their wider network of health service providers. This leads to unnecessary reordering of tests, incomplete medical histories, or medication information resulting in suboptimal patient care, occasionally endangering the patient's life [10].

Importantly, the BHS could serve as an education platform, "upskilling" patients by providing them with accurate information about their condition, treatment options, medications, and self-management strategies. In addition, if patients are prompted to record their clinical data (such as blood sugar levels or blood pressure), medication adherence, or side effects, these data could be aggregated and viewed by the patients and their treatment team to track the illness trajectory and management over time. Moreover, routine symptom assessment could trigger automatically dispatched symptom management advice in real time, suggest appropriate support services available locally, permit clinician notification when a patient's tracked symptoms indicated the need to review, or prompt the patients to contact their doctor or emergency services if the symptom is potentially life threatening.

The BHS could include digitally supported platforms or ecosystems that integrate care by linking not only electronic records but also patients' apps and devices with their health record, as well as providing virtual therapeutic spaces for clinicians and patients to interact. Examples of such systems include the Synergy platform, used in treatment of mental health and prevention of self-harm by identifying and rapidly responding to suicidal ideation among young people [21]. Through "upskilling" individuals, this initiative would empower people to have greater ownership over the management of their health. Finally, the BHS could be used to also provide clinicians with point-of-care decision support by providing diagnostic algorithms to be used in conjunction with clinical assessments and up-to-date, evidence-based, and easily accessible optimal care pathways suitable for their patient's condition, which will also serve as continuing professional development. Thus, although an EMR is the core of the BHS, there are a number of other equally important components required to implement the BHS.

The Current Landscape of Digital Health in Australia

Some components of an EMR system are already being used in Australia [22]. EMRs are commonly used in primary care, and they are being introduced into hospital settings [23,24]. Several large scale Australian hospitals have fully implemented an EMR and are using the system to electronically share prescription and medication details directly with pharmacy, eliminating written scripts, which avoids errors associated with translating handwriting [22]. This results in a safer and more efficient medication management [24]. Importantly, this system gives health care professionals more detailed access to information about medications taken, improving informed clinical decision making, which is particularly beneficial for individuals with comorbid conditions who are taking multiple medications [23,24]. These digital hospitals are also using EMRs to provide electronic orders to diagnostic providers, such as radiology and pathology within the hospital setting [23,24]. Some other areas where an EMR could provide significant benefits include improved safety and health care quality and greater patient, clinician, and administrator satisfaction, as well as cost savings and revenue gains, such as through reduced drug expenditure [25].

In Australia, this technology is being rapidly adopted by hospitals around the country (see Table 1), which presents some examples of technologies being adopted by Australian hospitals and health care practices.

Australia's "My Health Record" was launched in July 2012, and it is a patient-controlled, secure Web-based summary of health information, which can be accessed by individuals by using their computer or smartphones. This initiative will greatly assist Australia in achieving an integrated digital health system. Patients can control what clinical information is uploaded and select to share this information with their primary care physician, specialists, hospitals, and other health care providers. According to Australia's National Digital Health strategy, by 2022, the "My Health Record" system will allow clinicians to share patient information with other health care providers, using a nationally consistent, standards-based approach to secure messaging [30].

XSL•FO
RenderX

Table 1. Examples of digital health technologies adopted by Australian hospitals and health care practices.

Location	Digital health technology
South Australia	The state of South Australia has introduced an "Enterprise Patient Administration System," which is an integrated EMR[a] system for every patient admitted to a public hospital or health service, progressively being rolled out across all metropolitan public hospitals and a network of general practices in South Australia [26].
Tasmania	The Tasmanian state government has implemented a real-time reporting and recording system of controlled drugs. This system allows prescribers and pharmacists to determine what Schedule 8 drugs have been dispensed for a patient, if a patient has had his or her access to drugs of dependence restricted, or if a patient has been declared to be drug dependent or drug seeking by a medical practitioner [27].
Australia wide	Australia's "My Health Record" is a patient-controlled, secure Web-based summary of health information, which can be accessed by individuals by using their computer or smartphones.
The Northern Territory	The Northern Territory of Australia, which is geographically very remote and lacks many specialist health services, routinely provides telehealth services, which have increased appointment attendance and reduced patient travel time and expenses [28].
Queensland	The long-term strategy for the state of Queensland is a Digital Hospital and Health Service [29]; therefore, the government's next intended step is to link the EMR among care settings across the health system [29]. The Queensland state health department has already invested in the development of an integrated EMR across 9 facilities and 7 hospital and health services, with the aim of enabling the exchange of information among primary, community, and acute care settings across the health system [29]. In Queensland [30], all public hospitals and health services are already connected to the "My Health Record" system, and a number of hospitals and health services are connected in other parts of Australia [31].
Victoria	In the state of Victoria, several health services have implemented an electronic referral process, so that referrals are sent electronically from 1 health care service to another rather than via fax or post [32].

[a]EMR: electronic medical record.

Furthermore, individuals and health care providers will have access to information about prescribed and dispensed medicines, which will limit abuse of prescription medications, such as opioids [33,34]. Patients will be able to digitally request their medicine on the Web, and all pharmacists will have access to electronic prescribing [30,35]. The National Digital Health Strategy is also prioritizing an end-of-life care pilot test to explore how advanced care planning documents can be incorporated into the "My Health Record," to be more accessible to treating health professionals [30]. In addition, the "My Health Record" provides a data platform for new interfaces to be added onto the system, and a number of apps are currently authorized to connect the "My Health Record" system, including "Health Engine," which connects individuals to health practitioners and allows Web-based booking of appointments [36]. The Australian government has developed a Web-based mental health portal that provides information about mental health apps and services. "My Aged Care" is a similar portal where individuals can access information on aged care and related care services.

In the pilot scheme before 2018, about 20% of Australians registered for a "My Health Record" [37]. This scheme was mandated nationally from mid-2018, with a record being automatically created for all Australians, unless they "opted out" [38-41]. Currently, 90% of Australians have a "My Health Record" [42]. According to Australia's National Digital Health strategy, the "My Health Record" is set to be the core component of Australia's future national digital health service [30]. In Queensland, [30] all public hospitals and health services are already connected to the "My Health Record" system, and a number of hospitals and health services are connected in other parts of Australia [31]. In Australia alone, the potential economic benefit of transitioning to an EMR system in all public acute and private hospitals equates to approximately Aus $1.76 billion annually [25]. This would be far greater if expanded to community and primary care settings [25]. Telehealth is an important part of digital health and a BHS. Indeed, Mate and Salinas (2014) [19] highlighted that many elements of face-to-face clinic visits can be performed at home, with the aid of modern technology, which is consistent with and important for the development of a BHS, particularly for those individuals who lack access to specialist care because of geographical isolation or limited mobility [43]. It has been estimated that in Australia, telehealth can not only improve access to medical care and increase convenience for patients but it may also reduce cost by up to AU $3 billion annually through reduced residential care costs, emergency admissions, potentially preventable hospitalizations, and Royal Flying Doctors Services in rural areas, as well as patient transport and travel [44,45]. Currently, the Australian Federal and State Governments are expanding digital health to improve connectivity among service providers in community and primary care settings and reach people in their own homes [25]. Teleconference consultations between clinicians and individuals in remote areas and living in residential care facilities are currently being conducted as part of a home monitoring study and funded through the Australian Medicare system [44,45]. To facilitate telehealth services, a local telecommunications company, Telstra, has developed the "My Care Manager," a tablet-based platform where clinicians, individuals, and carers can video conference, where medication details, scheduled services, and test reminders and care plans can be stored, and where data from medical devices, such as body weight, glucose, and pulse oximeters can be uploaded and remotely monitored by clinicians [46]. If part of an integrated BHS, it is possible that additional medical requirements could be done remotely or more locally, for example, medication reviews could be done virtually [47], and individuals could have blood draws taken at local centers [48,49]. Currently, national agencies in Australia have tested a system to remotely monitor vital signs, such as electrocardiography, heart rate, spirometry, noninvasive blood pressure, oxygen saturation, body weight,

XSL•FO
RenderX

glucometry, and body temperature, in a cohort of elderly individuals with chronic illnesses [45,50]. Finally, a number of technology providers in Australia, such as Springday [51] or InnoWell [52] are providing integrated platforms that provide support for new models of care that integrate devices, apps, and virtual consultations into one-stop consumer-facing platforms in the management of chronic diseases, such as mental health or cancer [21]. These initiatives demonstrate the potential of an integrated BHS, which could integrate EMRs, remote access to medical care, and consultation and health tracking, among other functions. In summary, an integrated digital health system is evolving in Australia to allow individuals to access test results, set up appointments, medication reminders, test reminders, and have teleconsultations with health care providers, but routine implementation of these systems is patchy, and the systems that are in place are often not linked. It is yet to be articulated how this new model of health care would operate on a population level.

Bespoke Health Care System: Moving Electronic Medical Records From an Information Depository to a Patient-Driven Management Platform

We argue that an intelligent approach using electronic personal medical records to place individuals at the center of their care and "upskill" them to have ownership over the management of their health care will lead to greater efficiencies and better patient outcomes. Indeed, individuals have expressed a desire to be empowered. The Australian Digital Health Agency led the extensive "Your health. Your say" consultation to inform the development of the Australia's National Digital Health strategy. More than 3000 consumers, carers, health care providers, community groups, professional bodies, and other stakeholders attended 103 forums, workshops, webcasts, and town hall meetings across Australia, and over 1000 submissions and survey responses were analyzed for key themes [30]. The first identified theme was that individuals want to take control of their health care decisions and need access to their personal health information to do this [30]. Participants identified that they wanted to be able to manage their medication, request prescription refills, and track their health status by using their smartphone [30,35]. An integrated BHS would allow individuals to track symptoms, have access to diagnostic test results and treatment decision aids, and obtain reliable information about their diagnoses, management plans and medications, and self-management strategies for disease, symptom, or side-effect management. In addition, these systems would support the integration of apps and devices, health data, and virtual care to support new, potentially disruptive models of care. Previous research shows that it is possible to increase self-management of care through technology [53]. For example, an electronic health diary and symptom tracker/health monitoring tool have been shown to improve self-efficacy in individuals with acute coronary syndrome [54]. This proposal mimics the flipped classroom model, whereby individuals can engage in didactic learning at home [17,18,55], and their time with clinicians can therefore be focused on discussing and solving problems and applying the knowledge learned earlier [55]. The BHS could also be used to assist individuals to adhere to clinical interventions. Indeed, electronic prompts and reminder alerts have been shown to increase patient engagement with digital interventions [56]. Symptom monitoring and patient feedback through the BHS could be used to assign patients to the appropriate level of stepped care, and the provision of care tailored to the individual's level of need could be delivered using Web-based interventions through the system. Piette et al [57] have developed a personalized cognitive behavior therapy pain management service that adapts to each patient's unique and changing needs to automatically personalize the intensity and type of patient support, using feedback from patients about their progress. Technology would enable the tailoring to individual, clinical, and personal circumstances by using reinforcement learning algorithms to deliver information that is most relevant to each user, similar to those algorithms used by Netflix, Google, and Amazon [58]. This system could also involve and "upskill" family/carers, which would increase the capacity of the health care system by expanding nonprofessional community care. Research shows that an interactive Web portal providing targeted support for informal caregivers of persons with dementia and professionals improves the acquisition of information, interactions between carers and professionals, access to support from home, and self-perceived empowerment in health-related decisions [59]. A further advantage of the BHS is that it could be used to link interested individuals to peer support networks, which could result in greater social connectedness and support and improved coping strategies [60,61]. Individuals could "opt in" to be linked with other individuals with similar diagnoses and circumstances. Finally, this system would enable consumer ownership of the individuals' health data, increasing consumer health literacy, engaging consumers in treatment decision making, activating health promoting behaviors and self-management, and supporting consumers to remain in their homes andlocal communities. Such a system would place consumers in the "driver's seat" in regard to the management of their own care. The increasing focus on self-management of chronic disease in many respects places the clinician in the position of coach rather than director of health care, much in the same way education has shifted from a model of the all-knowing expert who lectures to students to a model focused on cocreation of learning, where the teacher guides rather than directs learning through a collaborative environment. Textbox 1 displays an imaginary case study of how the BHS may operate from the patient's perspective.

XSL•FO
RenderX

Textbox 1. Imaginary case study of how the Bespoke Health Care System may operate from the patient's perspective.

Georgia is in her late 70s and lives at home with her husband Frank and their 2 corgis, Lucy and Roxy. Georgia has diabetes, which is generally well controlled. Georgia's smartphone beeps at her and notifies her that she is due to meet with her General Practitioner, Susan, for a regular checkup. Georgia and her General Practitioner are both connected to a Bespoke health care platform that enables them to communicate with one another, as well as other health care professionals, through secure messaging and web-based face-to-face meetings. It is this platform that has notified Georgia that she is due to meet with Susan.

Using the Bespoke health care platform on her smartphone, Georgia sees that Susan has a free appointment time available on Thursday morning, which Georgia books. The Bespoke health care platform also notifies Georgia that she is due to have some routine blood tests as part of her checkup, and it shows her some available times at the pathology clinic around the corner from Georgia's house. There is a spot available tomorrow afternoon, and Georgia selects that time. She will take Lucy and Roxy with her on her walk to the pathology clinic.

After her appointment, the pathology clinic uploads Georgia's test results to her Bespoke health care platform. This allows Susan, Georgia's General Practitioner, to see her results instantly, and if appropriate, to make these available for Georgia to view as well. Fortunately, all of Georgia's tests results are normal, and Susan sends Georgia a message telling her this.

On Thursday, Georgia and Susan have a web-based face-to-face appointment using the Bespoke health care platform. The platform securely records their meetings. At any time, Georgia and Susan both go back and review anything that they discussed during their appointment.

After her appointment, Georgia realizes that she is running low on her medication. She logs in to her Bespoke health care platform and requests a refill of her regular medication, the details of which are already stored in her profile. Susan remotely approves this request, and it is automatically sent to the pharmacy of Georgia's choosing. Georgia selects the pharmacy closest to her house. The pharmacist then makes the medication ready for Georgia to pick up, and the pharmacist notifies her using the Bespoke health care platform.

On Saturday morning, Georgia and Frank decide to take Lucy and Roxy for walk to the pharmacy to collect Georgia's medication. They will also visit their favourite café for breakfast beforehand. Unfortunately, after collecting her medication, Georgia rolls her ankle and has a small fall. She is alright, but she needs to go to the hospital for some monitoring. At the hospital, Georgia's treating physician, Alex, is able to view Georgia's medical history, medication information, and treatment plan, as well as any allergies or other relevant medical information, using the Bespoke health care platform. All of the information about Georgia's visit to the hospital, including her treatment plan, is also uploaded to the Bespoke health care platform, and Georgia's General Practitioner, Susan, is notified to check in with Georgia regarding her ankle at their next appointment.

On Sunday afternoon, Georgia is settling in back at home, and The Bespoke health care platform automatically sends her some practical tips about caring for her ankle and a reminder to take the medication that she collected from the pharmacy yesterday, as per Susan's recommendations, and it's a good thing it did, as Georgia was preoccupied thinking about her ankle.

Necessary Conditions for Successful Adoption of Bespoke Health Care

The critical element to achieve a BHS is the establishment of a universal personalized EMR, which has already occurred with the "My Health Record [30]" in Australia. All individuals would need to have access to fast, reliable, ubiquitous internet, and there would need to be widespread ownership and proficient use of mobile devices. Community and specialist involvement and infrastructure would be necessary. A BHS would require the involvement of local community centers to provide education in the BHS, host peer support groups, deliver coaching in self-management, and possibly use motivational interviewing techniques to increase engagement with the system [62]. The system would need to be underpinned by best available evidence, regularly updated, and have high acceptability with health care providers, patients, carers, and privacy advocates. The BHS would also present new opportunities for digitally supported communication between health providers and patients to support innovative models of care. Critically, universal health records need to be able to securely store data from such emerging systems. Potential benefits of the BHS are multiple. First and foremost, the system would place the consumers at the center of their care and promote self-management. Individuals living in regional/remote areas would have greater access to peak health services and specialist clinicians via teleconsultations. The application of optimal care pathways could be offered, regardless of the patient's location, which would reduce regional variations in health outcomes [1]. Patients would have access to real-time, evidence-based self-management advice. There would be reduced duplication and errors. Moreover, clinical encounters would be optimized by automating routine clinical activities, such as assessment of symptoms or side effects. Furthermore, more clinical time could be devoted to health care decision making and education, such as reinforce health promoting behaviors. The intent is that the BHS would enable the routine and universal delivery of best-evidence health care, resulting in optimal health outcomes and better patient experiences of the health system. The system should demonstrate improved or equivalent patient outcomes, with a reduction or no increase in health care costs. Finally, iterative and quality improvement processes based on performance data could be built into the BHS. The development of this system requires a team approach. On a government level, there would need to be regulation and oversight, standards, systems, and infrastructure to support the BHS. Health care funders, such as Australia's Medicare and insurance agencies, would need to change the architecture of their funding models, which currently center around face-to-face medical consultations with an individual patient and the delivery of procedures. Progress is being made in this domain, with teleconsultations between clinicians and individuals in remote areas and Residential Care Facilities, which are currently being funded through the Medicare system [44,45]. Industries would need to develop and roll out new business and care models based around personalized medicine and informed by best-evidence practice. Universities can play a key role in training future health professionals' in digital health care systems, as well as supporting the analysis and interpretation of big data, research, evaluation, and iterative improvement of the BHS. The needs

XSL•FO
RenderX

of consumers and advocacy group end users would need to be identified and integrated, as was done in the "Your health. Your say" consultation [30]. The experiences of clinicians working within the health system must also be considered and integrated through processes, such as codesign models.

To be accepted and utilized by patients and health professionals, the initiative must be designed to be acceptable to the end users (both patients and health professionals), cater to individuals' unique needs to place minimal demands on the health system infrastructure, and be scalable and rapidly disseminated into usual care if successful. Co-design is crucial to the intervention development process to design for the user. Co-design seeks to understand the lived experience of end users, making the everyday practices and contexts of the target audience important resources to inform the design of the intervention [63]. Co-design refers to collective creativity as it is applied across the whole span of a design process [64]. Specifically, this entails experts (researchers, designers, or developers) working together with end users (consumers) from concept creation, prototype review, to final product. In addition, theoretical frameworks should be considered to provide the basis to optimize implementation strategies and to identify implementation barriers, strategies to address these barriers and methods to explore mediating mechanisms [65]. A comprehensive understanding of the barriers and enablers to implementation of a BHS in clinical care is necessary to support the development of effective implementation strategies [66,67]. Dansereau and colleagues outlined that there are 4 levels of analysis relevant to organizational behavior, these are persons, dyads, groups, and organizations [68-70]. Evidence shows implementation requires whole system change, involving both the individual and organization [71]. Theoretical frameworks can help explain why implementation efforts succeed or fail [72], but the biggest challenge influencing the implementation of new interventions is behavior change, with many frameworks therefore focusing on human behavior theories [73,74]. The Theoretical Domains Framework (TDF) developed by Michie and colleagues [74] used expert consensus to identify 12 domains of behavior determinants that could be used in implementation research. Utilizing frameworks similar to TDF will be useful in identifying the individual- and organizational-level barriers most relevant to the implementation of a BHS. Consumer ownership of data and privacy considerations also need to be addressed. The Australian Digital Health Agency acknowledged this concern in the National Digital Health strategy, and it has established the Digital Health Cyber Security Centre. Its primary purpose is to protect the national digital health systems and personal health information of Australians from cyber threat and raise the security posture of the Australian health sector. The Digital Health Cyber Security Centre partners with national and international cyber security organizations, across the government and private sector, to improve knowledge of the cyber threat and leverage shared expertise and material across organizations [30]. The recent funding of the Digital Health Cooperative Research Centre by the Australian Federal Government that links academia with services, government, and industry also provides a vehicle for the development and support of new national, coordinated approaches to the development of new technology-driven models of care [75].

Conclusions

The ultimate objectives of the proposed BHS are reducing pressure on the current health care system, delivering best-practice health care, regardless of location, and placing consumers in a position to direct their own health care. Delivery of personalized, evidence-based health care, regardless of location, is likely to improve disease outcomes, quality of life, and patient experiences. "Upskilling" individuals and enabling easy information sharing can improve communication between individuals and health care providers. The BHS can offer real-time, accurate data collection and information dissemination. Importantly, delivery of best practice health care in regional locations would aid in decentralizing health services by stimulating health industry growth in regional centers, as well as reducing the pressure and resource constraints in urban hospitals. This system would also contribute to the development of a next generation workforce, which is responsive to emerging challenges, particularly in the digital domain. By proposing BHS, we offer an intelligent model of health care that is ubiquitous, efficient, and continuously improving. The system would allow research and iterative improvements to be embedded into health systems. Finally, reducing service production and delivery waste is likely to result in economic benefits.

Conflicts of Interest

None declared.

References

1. Australian Institute of Health and Welfare. Canberra: AIHW; 2012. Risk factors contributing to chronic disease URL: https://www.aihw.gov.au/getmedia/74121d1b-69ca-4a34-a08a-51b741ea26b2/12724.pdf.aspx?inline=true [accessed 2019-05-14] [WebCite Cache ID 78NJ9lcLS]
2. United Nations. 2015. World Population Ageing 1950-2050 URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf [accessed 2019-05-14] [WebCite Cache ID 78NJXaDIi]
3. United Nations. 2015. World Population Ageing 2015 Report URL: https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2015_Report.pdf [accessed 2019-05-14] [WebCite Cache ID 78NJlThOk]
4. World Health Organization. 2002. The world health report 2002, reducing risks, Promoting Healthy Life URL: https://www.who.int/whr/2002/en/whr02_en.pdf?ua=1 [accessed 2019-05-15] [WebCite Cache ID 78Oz6ZVKQ]

XSL•FO
RenderX

5. World Health Organization. 2017. World health statistics 2017: monitoring health for the SDGs, Sustainable Development Goals URL: https://www.who.int/gho/publications/world_health_statistics/2017/en/ [accessed 2019-05-14] [WebCite Cache ID 78NK8lYMU]
6. Australian Institute of Health and Welfare. Canberra: AIHW; 2016. 25 years of health expenditure in Australia 1989–90 to 2013–14 URL: https://www.aihw.gov.au/getmedia/bcfd0546-98d4-4111-bbfa-da46b8e583a8/18011.pdf.aspx?inline=true [accessed 2019-05-14] [WebCite Cache ID 78NKyMjjo]
7. Australian Institute of Health and Welfare. 2016. 25 years of health expenditure in Australia 1989–90 to 2013–14 URL: https://www.aihw.gov.au/getmedia/bcfd0546-98d4-4111-bbfa-da46b8e583a8/18011.pdf.aspx?inline=true [accessed 2019-05-14] [WebCite Cache ID 78NLBt18s]
8. Roughead L, Semple S, Rosenfeld E. Australian Commission on Safety and Quality in Health Care. Sydney: ACSQHC; 2013. Literature Review: Medication Safety in Australia URL: https://safetyandquality.gov.au/wp-content/uploads/2014/02/Literature-Review-Medication-Safety-in-Australia-2013.pdf [accessed 2019-05-14] [WebCite Cache ID 78NM1Xj6c]
9. Seeniyar R. Pharmaceutical Benefits Scheme (PBS). Canberra: Department of Health; 2016. Expenditure and Prescriptions Twelve Months to 30 June 2017 URL: http://www.pbs.gov.au/statistics/expenditure-prescriptions/2016-2017/expenditure-and-prescriptions-twelve-months-to-30-june-2017.pdf [accessed 2019-05-14] [WebCite Cache ID 78NMSUayx]
10. Bentley TG, Effros RM, Palar K, Keeler EB. Waste in the U.S. Health care system: a conceptual framework. Milbank Q 2008 Dec;86(4):629-659 [FREE Full text] [doi: 10.1111/j.1468-0009.2008.00537.x] [Medline: 19120983]
11. Productivity Commission. 2005. Productivity Commission Research Report URL: https://www.pc.gov.au/inquiries/completed/health-workforce/report [accessed 2019-05-14] [WebCite Cache ID 78NNKGR2K]
12. Taylor M, Hill S. Australian Healthcare and Hospital Association. 2014. Consumer expectations and healthcare in Australia URL: https://ahha.asn.au/sites/default/files/docs/policy-issue/deeble_issues_brief_nlcg-3_consumer_expectations_and_healthcare_in_australia.pdf [accessed 2019-05-14] [WebCite Cache ID 78NOD2hn0]
13. Beyond Blue. 2013. National mental health survey of doctors and medical students URL: https://www.beyondblue.org.au/docs/default-source/research-project-files/bl1132-report---nmhdmss-full-report_web [accessed 2019-05-14] [WebCite Cache ID 78NOb94oe]
14. Shanafelt T, Dyrbye L. Oncologist burnout: causes, consequences, and responses. J Clin Oncol 2012 Apr 10;30(11):1235-1241. [doi: 10.1200/JCO.2011.39.7380] [Medline: 22412138]
15. Shaw T, McGregor D, Brunner M, Keep M, Janssen A, Barnet S. What is eHealth (6)? Development of a conceptual model for eHealth: qualitative study with key informants. J Med Internet Res 2017 Dec 24;19(10):e324-e341 [FREE Full text] [doi: 10.2196/jmir.8106] [Medline: 29066429]
16. CSIRO. 2018. Future of Health: Shifting Australia's focus from illness treatment to health and wellbeing management URL: https://www.csiro.au/en/Do-business/Futures/Reports/Future-of-Health [accessed 2019-05-14] [WebCite Cache ID 78NP0wRRH]
17. Lage J, Platt J, Treglia M. Inverting the classroom: a gateway to creating an inclusive learning environment. J Econ Educ 2000;31(1):30. [doi: 10.2307/1183338]
18. Jonathan B. Flip Your Classroom: Reach Every Student in Every Class Every Day. USA: International Society For Technology In Education; 2019.
19. Mate KS, Salinas G. Flipping primary health care: a personal story. Healthc (Amst) 2014 Dec;2(4):280-283. [doi: 10.1016/j.hjdsi.2014.10.003] [Medline: 26250637]
20. Australian Digital Health Agency. 2012. National eHealth Transition Authority, 2012 Annual Report URL: https://www.digitalhealth.gov.au/about-the-agency/publications/reports/annual-report [accessed 2019-05-14] [WebCite Cache ID 78NQ3hPhe]
21. Iorfino F, Davenport TA, Ospina-Pinillos L, Hermens DF, Cross S, Burns J, et al. Using new and emerging technologies to identify and respond to suicidality among help-seeking young people: a cross-sectional study. J Med Internet Res 2017 Dec 12;19(7):e247 [FREE Full text] [doi: 10.2196/jmir.7897] [Medline: 28701290]
22. Mitchell C. InSight. 2016. Digital revolution rolling out in hospitals URL: https://insightplus.mja.com.au/2016/42/digital-revolution-rolling-out-in-hospitals/ [accessed 2019-05-14] [WebCite Cache ID 78NQxB3e6]
23. Sullivan C, Staib A, Ayre S, Daly M, Collins R, Draheim M, et al. Pioneering digital disruption: Australia's first integrated digital tertiary hospital. Med J Aust 2016 Nov 7;205(9):386-389 [FREE Full text] [Medline: 27809727]
24. Swannell C. Medical Journal of Australia. 2013. Fully wired hospital for Hervey Bay URL: https://www.mja.com.au/journal/2013/199/4/fully-wired-hospital-hervey-bay [accessed 2019-05-14] [WebCite Cache ID 78NR8b7cv]
25. Forsythe J, Rogan C, Dimkin D, Strain R, Curran J, Odhav V. PwC Australia. 2016. Australia can see further by standing on the shoulders of giants - Driving digital transformation by adopting Meaningful Use legislation URL: https://www.pwc.com.au/publications/pdf/digital-hospital-2016.pdf [accessed 2019-05-14] [WebCite Cache ID 78NRElQdM]
26. Department for Health and Ageing, Government of South Australia. 2017. Enterprise Patient Administration System (EPAS) URL: https://www.sahealth.sa.gov.au/wps/wcm/connect/9489b046-246c-4d71-a0a0-012e233e3dac/EPASReview-Issues%26DirectionsConsultationPaper+v1.0.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE-9489b046-246c-4d71-a0a0-012e233e3dac-mFHYcDe

XSL•FO
RenderX

27. Department of Health and Human Services, Tasmanian Government. DORA - Drugs and Poisons Information System Online Remote Access URL: https://www.dhhs.tas.gov.au/psbtas/publications/general/dora [accessed 2019-05-14] [WebCite Cache ID 78NRfih2m]
28. Aboriginal Medical Services Alliance Northern Territory, Department of Health. Evaluation of the PATS-Telehealth Project URL: http://www.amsant.org.au/wp-content/uploads/2016/03/PATS-Telehealth-Report-FINAL.pdf [accessed 2019-05-14] [WebCite Cache ID 78NRjaQTB]
29. Queensland Health. 2015. 21st Century Healthcare eHealth Investment Strategy URL: https://www.health.qld.gov.au/__data/assets/pdf_file/0031/442939/ehealthinvestmentstrategy.pdf [accessed 2019-05-14] [WebCite Cache ID 78NRtE4cd]
30. Australian Digital Health Agency. Australia's National Digital Health Strategy URL: https://conversation.digitalhealth.gov.au/australias-national-digital-health-strategy [accessed 2019-05-15] [WebCite Cache ID 78NQ9CDJV]
31. My Health Record. 2017. Private hospitals connected to the My Health Record system URL: https://www.myhealthrecord.gov.au/about/who-is-using-digital-health/private-hospitals-connected-my-health-record-system [accessed 2019-05-14] [WebCite Cache ID 78NSCE5ZH]
32. Premier of Victoria. Victoria: Victorian Governement; 2016. GP e-Referrals Rolling Out In Melbourne's East For Better Patient Care URL: https://www.premier.vic.gov.au/gp-e-referrals-rolling-out-in-melbournes-east-for-better-patient-care/ [accessed 2019-05-14] [WebCite Cache ID 78NS3eJe1]
33. Rasubala L, Pernapati L, Velasquez X, Burk J, Ren Y. Impact of a mandatory prescription drug monitoring program on prescription of opioid analgesics by dentists. PLoS One 2015;10(8):e0135957 [FREE Full text] [doi: 10.1371/journal.pone.0135957] [Medline: 26274819]
34. Franklin G, Sabel J, Jones CM, Mai J, Baumgartner C, Banta-Green CJ, et al. A comprehensive approach to address the prescription opioid epidemic in Washington State: milestones and lessons learned. Am J Public Health 2015 Mar;105(3):463-469. [doi: 10.2105/AJPH.2014.302367] [Medline: 25602880]
35. Analysis and Policy Observatory. 2017. Your Health Your Say Consultation Sydney URL: https://apo.org.au/sites/default/files/resource-files/2016/11/apo-nid72360-1110941.pdf [accessed 2019-05-14] [WebCite Cache ID 78NQP4pog]
36. My Health Record. Canberra: Australian Digital Health Agency; 2017. View your record using an app URL: https://www.myhealthrecord.gov.au/for-you-your-family/howtos/view-your-record-using-app [accessed 2019-05-14] [WebCite Cache ID 78NSLfbfI]
37. My Health Record. My Health Record statistics URL: https://www.myhealthrecord.gov.au/statistics [accessed 2019-05-14] [WebCite Cache ID 78NSRejUN]
38. My Health Record. Canberra: Australian Digital Health Agency; 2017. URL: https://www.myhealthrecord.gov.au/ [accessed 2019-05-14] [WebCite Cache ID 78NT65kYV]
39. Walsh L, Hill S, Allan M, Balandin S, Georgiou A, Higgins I, et al. A content analysis of the consumer-facing online information about My Health Record: implications for increasing knowledge and awareness to facilitate uptake and use. Health Inf Manag 2018 Sep;47(3):106-115. [doi: 10.1177/1833358317712200] [Medline: 28589741]
40. Hemsley B, Georgiou A, Carter R, Hill S, Higgins I, van Vliet P, et al. Use of the My Health Record by people with communication disability in Australia: a review to inform the design and direction of future research. Health Inf Manag 2016 Dec;45(3):107-115. [doi: 10.1177/1833358316652060] [Medline: 27269277]
41. Skilton N. My Health Record to deliver "new class" of data. Public Health Res Pract 2016 Jan 28;26(1):e2611611 [FREE Full text] [doi: 10.17061/phrp2611611] [Medline: 26863174]
42. Australian Digital Health Agency. 2019. 9 out of 10 Australians have a My Health Record URL: https://www.digitalhealth.gov.au/news-and-events/news/9-out-of-10-australians-to-have-a-my-health-record [accessed 2019-05-14] [WebCite Cache ID 78NStExcW]
43. Bradford N, Caffery L, Smith A. Telehealth services in rural and remote Australia: a systematic review of models of care and factors influencing success and sustainability. Rural Remote Health 2016;16(4):4268 [FREE Full text] [Medline: 27817199]
44. Celler BG, Sparks R, Nepal S, Alem L, Varnfield M, Li J, et al. Design of a multi-site multi-state clinical trial of home monitoring of chronic disease in the community in Australia. BMC Public Health 2014 Dec 15;14:1270 [FREE Full text] [doi: 10.1186/1471-2458-14-1270] [Medline: 25511206]
45. Celler B, Varnfield M, Nepal S, Sparks R, Li J, Jayasena R. Impact of at-home telemonitoring on health services expenditure and hospital admissions in patients with chronic conditions: before and after control Intervention Analysis. JMIR Med Inform 2017 Sep 8;5(3):e29 [FREE Full text] [doi: 10.2196/medinform.7308] [Medline: 28887294]
46. Telstra Corporation Limited. 2015. My care manager URL: https://www.telstrahealth.com/home/products/mycaremanager.html [accessed 2019-05-14] [WebCite Cache ID 78NVKp9cc]
47. Newsome J. Virtual medication therapy management: the cornerstone to a pharmacy renaissance. J Am Pharm Assoc (2003) 2013;53(2):159-162. [doi: 10.1331/JAPhA.2013.12188] [Medline: 23571623]
48. Leff B, Burton JR. The future history of home care and physician house calls in the United States. J Gerontol A Biol Sci Med Sci 2001 Oct;56(10):M603-M608. [doi: 10.1093/gerona/56.10.m603] [Medline: 11584032]

XSL•FO
RenderX

49. Heath G, Cameron E, Cummins C, Greenfield S, Pattison H, Kelly D, et al. Paediatric 'care closer to home': stake-holder views and barriers to implementation. Health Place 2012 Sep;18(5):1068-1073. [doi: 10.1016/j.healthplace.2012.05.003] [Medline: 22687749]
50. CSIRO. Home Monitoring of Chronic Diseases URL: https://www.csiro.au/en/Research/BF/Areas/Digital-health/Improving-access/Home-monitoring [accessed 2019-05-14] [WebCite Cache ID 78NTLtACB]
51. Springday. URL: https://corporate.myspringday.com.au/ [accessed 2019-05-16] [WebCite Cache ID 78P76GPmT]
52. InnoWell. URL: https://www.innowell.org/ [accessed 2019-05-16] [WebCite Cache ID 78P7YtCNO]
53. Peeters JM, Wiegers TA, Friele RD. How technology in care at home affects patient self-care and self-management: a scoping review. Int J Environ Res Public Health 2013 Oct 29;10(11):5541-5564 [FREE Full text] [doi: 10.3390/ijerph10115541] [Medline: 24173139]
54. Wolf A, Fors A, Ulin K, Thorn J, Swedberg K, Ekman I. An eHealth diary and symptom-tracking tool combined with person-centered care for improving self-efficacy after a diagnosis of acute coronary syndrome: a substudy of a randomized controlled trial. J Med Internet Res 2016 Feb 23;18(2):e40 [FREE Full text] [doi: 10.2196/jmir.4890] [Medline: 26907584]
55. Fulton K. ERIC - Education Resources Information Center. 2012. Upside Down and Inside Out: Flip Your Classroom to Improve Student Learning URL: https://eric.ed.gov/?id=EJ982840 [accessed 2019-05-14] [WebCite Cache ID 78NTvU09N]
56. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 8;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
57. Piette JD, Krein SL, Striplin D, Marinec N, Kerns RD, Farris KB, et al. Patient-centered pain care using artificial intelligence and mobile health tools: protocol for a randomized study funded by the US Department of Veterans Affairs Health Services Research and Development Program. JMIR Res Protoc 2016 Apr 7;5(2):e53 [FREE Full text] [doi: 10.2196/resprot.4995] [Medline: 27056770]
58. Lihong L, Chu W, Langford J, Schapire R. A contextual bandit approach to personalized news article recommendation. In: Proceedings of the 19th international conference on World wide web. 2010 Presented at: WWW'10; April 26-30, 2010; Raleigh, North Carolina, USA p. 661-670.
59. Schaller S, Marinova-Schmidt V, Setzer M, Kondylakis H, Griebel L, Sedlmayr M, et al. Usefulness of a tailored ehealth service for informal caregivers and professionals in the dementia treatment and care setting: the eHealthMonitor Dementia portal. JMIR Res Protoc 2016 Apr 5;5(2):e47 [FREE Full text] [doi: 10.2196/resprot.4354] [Medline: 27050401]
60. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
61. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci 2016 Apr;25(2):113-122 [FREE Full text] [doi: 10.1017/S2045796015001067] [Medline: 26744309]
62. Schneider H, Hill S, Blandford A. Patients know best: qualitative study on how families use patient-controlled personal health records. J Med Internet Res 2016 Feb 24;18(2):e43 [FREE Full text] [doi: 10.2196/jmir.4652] [Medline: 26912201]
63. Bate P, Robert G. Bringing User Experience to Healthcare Improvement: The Concepts, Methods and Practices of Experience-based Design. Oxford: Radcliffe Publishing; 2007.
64. Steen M, Manschot M, de Koning N. Benefits of co-design in service design projects. Int J Des 2011;5(2):53-60 [FREE Full text]
65. Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. BMJ Qual Saf 2015 Mar;24(3):228-238 [FREE Full text] [doi: 10.1136/bmjqs-2014-003627] [Medline: 25616279]
66. Albers B, Pattuwage L. Centre for Evidence and Implementation. Australia: Evidence for Learning; 2017. Implementation in education - Findings from a Scoping Review URL: https://www.ceiglobal.org/application/files/2514/9793/4848/Albers-and-Pattuwage-2017-Implementation-in-Education.pdf [accessed 2019-05-14] [WebCite Cache ID 78NUXZ37S]
67. Domitrovich CE, Bradshaw CP, Poduska JM, Hoagwood K, Buckley JA, Olin S, et al. Maximizing the implementation quality of evidence-based preventive interventions in schools: a conceptual framework. Adv Sch Ment Health Promot 2008 Jul;1(3):6-28 [FREE Full text] [Medline: 27182282]
68. Dansereau F, Yammarino FJ, Alutto JA. Theory Testing in Organizational Behavior: The Varient Approach. Englewood Cliffs: Prentice-Hall; 1984.
69. Dionne SD, Gupta A, Sotak KL, Shirreffs KA, Serban A, Hao C, et al. A 25-year perspective on levels of analysis in leadership research. Leadersh Q 2014 Feb;25(1):6-35. [doi: 10.1016/j.leaqua.2013.11.002]
70. Yammarino FJ, Dionne SD, Uk Chun J, Dansereau F. Leadership and levels of analysis: a state-of-the-science review. Leadersh Q 2005 Dec;16(6):879-919. [doi: 10.1016/j.leaqua.2005.09.002]
71. Australian Centre for Health Services Innovation. 2016. Implementation Grant Funding Policy URL: http://www.aushsi.org.au/wp-content/uploads/2016/08/AusHSI-Implementation-Grant-Funding-Policy-V3-2016.pdf [accessed 2019-05-14] [WebCite Cache ID 78NUoEK5C]
72. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]

XSL•FO
RenderX

73. Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleesschauwer B, et al. Correction: World Health Organization estimates of the global and regional disease burden of 22 foodborne bacterial, protozoal, and viral diseases, 2010: a data synthesis. PLoS Med 2015 Dec;12(12):e1001940 [FREE Full text] [doi: 10.1371/journal.pmed.1001940] [Medline: 26701262]
74. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
75. Digital Health CRC. URL: https://www.digitalhealthcrc.com/ [accessed 2019-05-15] [WebCite Cache ID 78NUwZcBt]

Abbreviations

BHS: Bespoke Health Care System
EMR: electronic medical record
TDF: Theoretical Domains Framework

Edited by G Eysenbach; submitted 02.10.18; peer-reviewed by H Kondylakis, E Lattie; comments to author 05.03.19; revised version received 04.04.19; accepted 21.04.19; published 04.06.19.

Please cite as:
Schofield P, Shaw T, Pascoe M
Toward Comprehensive Patient-Centric Care by Integrating Digital Health Technology With Direct Clinical Contact in Australia
J Med Internet Res 2019;21(6):e12382
URL: https://www.jmir.org/2019/6/e12382/
doi:10.2196/12382
PMID:31165713

©Penelope Schofield, Tim Shaw, Michaela Pascoe. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Key Issues in the Development of an Evidence-Based Stratified Surgical Patient Safety Improvement Information System: Experience From a Multicenter Surgical Safety Program

Xiaochu Yu[1*], MD; Wei Han[2*], PhD; Jingmei Jiang[2*], PhD; Yipeng Wang[1], MD; Shijie Xin[3], MD; Shizheng Wu[4], MD; Hong Sun[5], MD; Zixing Wang[2], MSc; Yupei Zhao[1], MD

[1]Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, China

[2]Department of Epidemiology and Biostatistics, Institute of Basic Medicine Sciences, Chinese Academy of Medical Science, Beijing, China

[3]The First Hospital of China Medical University, Shenyang, China

[4]Qinghai Provincial People's Hospital, Xining, China

[5]Xiangya Hospital, Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Yupei Zhao, MD
Peking Union Medical College Hospital, Chinese Academy of Medical Sciences
No 1 Shuaifuyuan Wangfujing Dongcheng District
Beijing, 100730
China
Phone: 86 010 69155789
Email: zhao8028@263.net

Abstract

Surgery is still far from being completely safe and reliable. Surgical safety has, therefore, been the focus of considerable attention over the last few decades, and there are a growing number of national drives to improve it. There are also a number of large surgical complication reporting systems and system-based interventions, both of which have made remarkable progress in the past two decades. These systems, however, have either mainly focused on reporting complications and played a limited role in guiding practice or have provided nonselective interventions to all patients, perhaps imposing unnecessary burdens on frontline medical staff. We have, therefore, developed an evidence-based stratified surgical safety information system based on a multicenter surgical safety improvement program. This study discusses some critical issues in the process of developing this information system, including (1) decisions about data gathering, (2) establishing and sharing knowledge, (3) developing functions for the system, (4) system implementation, and (5) evaluation and continuous improvement. Using examples drawn from the surgical safety improvement program, we have shown how this type of system can be fitted into day-to-day clinical practice and how it can guide medical practice by incorporating inherent patient-related risk and providing tailored interventions for patients with different levels of risk. We concluded that multidisciplinary collaboration, involving experts in health care (including senior staff in surgery, nursing, and anesthesia), data science, health care management, and health information technology, can help build an evidence-based stratified surgical patient safety improvement system. This can provide an information-intensified surgical safety learning platform and, therefore, benefit surgical patients by delivering tailored interventions and an integrated workflow.

(J Med Internet Res 2019;21(6):e13576) doi:10.2196/13576

KEYWORDS

surgery; patient safety; information system; risk factors; evidence-based practice

Introduction

Globally, each year, more than 230 million operations are performed and at least 7 million patients develop significant surgical complications, including 1 million perioperative deaths (at least half of which are preventable) [1,2]. Preventing harm to patients and improving the safety of surgical patients has, therefore, drawn considerable attention over the last few decades, and there are growing national drives to improve surgical patients' safety [3]. There are also abundant opportunities for informatics-based improvements in perioperative care linked to the rapid development of

XSL•FO
RenderX

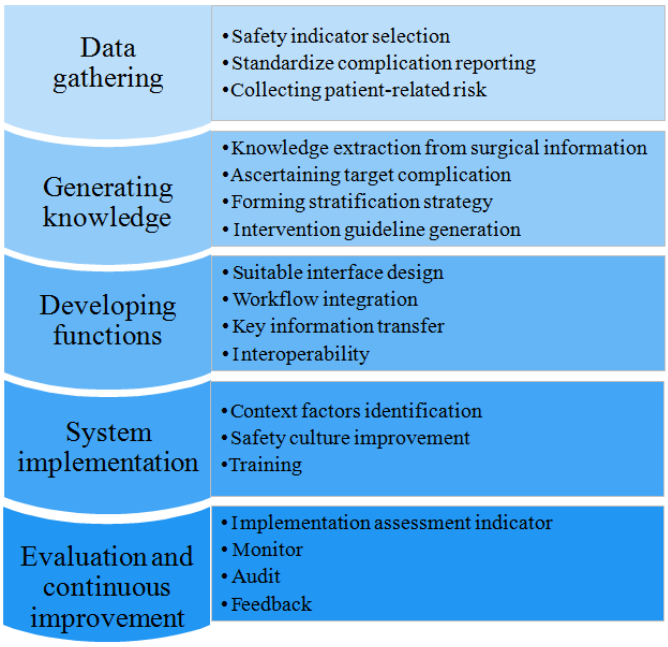
information technology use in medical care. Information systems, such as computerized physician order entry, automated dispensing, barcode medication administration, electronic medication reconciliation, and personal health records, are playing an increasingly important role in enhancing patient safety by reducing medication errors [4,5]. Up to 50.2% (470/936) of medical errors can be avoided through the use of information systems [6]. However, most of these systems are not designed specifically for surgery. Their original intention was to regulate clinicians' daily practice and allow health care providers to carry out routine jobs effectively, preventing potential errors [5]. Patient safety outcomes, which are crucial indicators for evaluating and improving surgical safety practice, cannot usually be obtained reliably through these systems. Extensive and carefully planned specialized information platforms for surgical sectors are, therefore, needed to collect data on patient safety [7]. Several nationwide surgical complication reporting and learning systems have, therefore, been developed in the past two decades, such as the United Kingdom's National Reporting and Learning System [8]. This was established in late 2003 as a voluntary scheme for reporting patient safety incidents, to support learning from these incidents. Another large-scale Web-based information platform is the American College of Surgeons National Surgical Quality Improvement Program, which dates back to the 1980s and now incorporates hundreds of hospitals across the United States. It was developed to gauge the quality of surgical programs across different hospitals. The primary function of both of these systems is surgical incident reporting. Measuring incidence alone, however, is not enough to guide routine clinical safety behavior and enhance safety. A number of system-based interventions have, therefore, also been developed, mainly focusing on regulating clinical behavior and following the publication of *To Err is Human* [9]. These included the Surgical Patient Safety System (SURPASS) checklist, which requires 11 forms (nearly 100 items) to be completed and documented by providers for each individual undergoing surgery [10,11]. These intervention strategies regulate the daily clinical practices of health care staff and thus improve patient safety, but they impose a heavy workload in complex clinical settings. This could increase fatigue and undermine the adoption of and compliance with these systems by frontline health care staff [12]. The number of successful system implementations is, therefore, relatively small, with conflicting findings on their effect on patient safety. This tends to lead to skepticism about the true effectiveness of these systems and emphasizes the necessity of developing a patient safety system with high implementation efficiency and low operational complexity [13,14].

Surgical operations are complex procedures. The perioperative care process is a unique and challenging environment that requires close collaboration among surgeons, anesthetists, and nurses. Medical staff can also encounter sophisticated patient pathophysiological conditions. The length of patients' stay in the hospital is relatively short in surgical departments, which has posed a significant challenge for surgeons in capturing the key factors influencing surgical outcomes and the timely transfer of key safety information to other members of the surgical team [15]. An information system capable of extracting knowledge from high volume and multi-sourced clinical data and supporting decisions in routine clinical operations could, therefore, improve efficiency and effectiveness. It could also provide a high degree of interoperability as well as information support, process management, and optimization in delivering evidence-based surgical safety interventions [16,17]. The complexities of the perioperative environment, however, can complicate the process of deployment and make technology implementation challenging. Some common issues in this environment must be addressed for successful deployment of information technology [18]. The development of this type of evidence-based patient safety information system (EPSIS), therefore, requires a holistic view. It needs to bring together clinical professionals, health care administrators, data scientists, and information technology engineers. This process can serve as an important catalyst in fostering a safety culture among frontline health workers, constructing a surgical safety ecosystem supported by data scientists, engineers, and administrators [19].

In 2014, a national project called Modern Surgery and Anesthesia Safety Management System Construction and Promotion (MSCP) was conducted in China. It aimed to improve perioperative patient safety. On the basis of this study, a perioperative surgical safety management information system was developed. This integrated patient data collection, processing, storage, and dissemination to support decision making, work control and documentation, and visualization [20]. To identify key elements and critical issues and formulate a framework to design, develop, and implement an EPSIS, a multidisciplinary panel of experts was assembled during the project period. This panel consisted of 10 medical experts, 5 nurses, 5 medical administrators, 3 data scientists, and 7 computer science engineers. All 30 experts attended several rounds of face-to-face consensus meetings, and widespread suggestions were collected from both the literature and panel members. Notes from project process meetings held by the central project group and the project executive groups from participating hospitals were also reviewed to identify practical challenges faced during the development of the project. These were refined and distributed to the panel members. Key methodological and implementation issues in designing and developing an evidence-based surgical safety information system were discussed, and recommendations on these issues were collected through these meetings. Finally, a summary of the framework was drafted and was circulated to panel members via email. Comments were collected until the group had reached a consensus. This paper discusses these issues (see Figure 1) in detail and uses practical examples from MSCP to illustrate them.

XSL•FO
RenderX

Figure 1. Critical issues in the life-cycle for designing and implementing an evidence-based surgical safety information system.



Developing and Gathering Data

Clinical data are usually complex and highly distributed. This poses challenges in collecting data to support the development of an EPSIS, because this type of system relies heavily on the management of high-quality surgical safety data. Using high-quality data also supports workflow management, monitoring, and evaluation of the surgical patient safety system. To satisfy the needs of the whole information system, the accessibility, reliability, and timeliness of data should be ensured.

Which Surgical Safety Indicators Should Be Collected?

Surgical safety promotion usually requires integrated interventions involving changes in a set of activities. These have long causal pathways and involve many factors that can influence the causal chain [21]. Recommended surgical safety indicators, therefore, include surgical complications and death and length of stays, which directly measure the observable harm. These indicators, unlike process or surrogate safety indicators (eg, error or culture), are identifiable and quantifiable and are more appropriate as natural endpoints in the *story* of patient safety [22]. The collection of data on these observable harms provides an easy way to investigate the causal chain. It could, therefore, help support learning about critical surgical safety issues linked to specific contexts as well as developing intuitive and target-sensitive functions for the safety improvement system [22-24].

How Should Complication Reporting Be Standardized?

Information about complications is not always readily available. A complication reporting procedure requires accepted principles of accrual, display, and analysis of complication data to be predefined to capture complications in a structural way [22,25,26]. This will allow meaningful comparisons of the incidence of reported complications across different hospitals or different periods within the same hospital [27]. The subsequent data processing could also benefit from this structural reporting. Several classification criteria had been proposed, of which the Clavien-Dindo classification is the most widely used. However, this type of classification system provides limited reference in standardizing complication reporting because it mainly focuses on ranking complication categories in an objective and reproducible manner, on the basis of the therapy used to correct them [28]. In 2002, Martin et al proposed 10 criteria that should be met when reporting complications following surgery [29]. These proposals could serve as a reference in establishing the criteria for reporting information about complications.

How Can We Retrieve and Integrate Information About Inherent Patient-Related Risk?

An incomplete data inventory leads to incomplete analyses. Electronic health records allow collection of particular elements of health-related information (eg, obesity, coronary heart disease, and hypertension) that are potentially associated with safety outcomes. However, it remains a challenge to form a multidisciplinary patient safety reference database because it needs to identify and track all data sources [30]. Knowledge of health data attributes, including data definitions, value sets, and other clinical coded content, is required. This means that the data retrieve process needs to involve information technology engineers, medical experts, and data scientists. Once the data have been captured, they can be filtered to support further clinical decision making by task and individual end user requirements. An additional step of data verification, introduced for data quality control, is also necessary. In MSCP, to ensure a trade-off between completeness and efficiency in information collection, a specialist panel was formed including medical experts, information engineers, and data scientists. By drawing on the literature and collecting expert opinions, this panel identified crucial information for inherent patient-related risk and determined how to obtain this information accurately and efficiently. To reduce user workload and transcription errors, a

XSL•FO
RenderX

data extraction strategy was established by the clinical experts and information system engineers for data readily obtainable from hospital information systems. For information that is not routinely collected by hospital information systems, or which requires a special reporting mechanism because of its importance (such as complications), a stand-alone electronic data capture system was developed. This included 3 separate subsystems (ward, intensive care unit, and operation room). Patient information was entered by an established data entry team once the patients had been admitted. Complications (using clear definitions) were entered within a week of the patient's discharge. The data collected from these 3 subsystems were centrally managed. Regular data quality audits were also conducted, and the results were reported monthly.

Generating and Sharing Knowledge

Information alone is not enough to improve safety. Knowledge about the spectrum of complications and potential risk information extracted from the data is essential to formulate guidelines to help decision making about which patients to prioritize and what measures to take to prevent surgical complications. Ultimately, these data-driven decisions could play an auxiliary role in supporting clinical decisions and lead to more effective and appropriate use of resources through better procedures. New data obtained through the system are likely to stimulate the updating of knowledge and, therefore, further improve decisions.

How Should We Prioritize Complications to Target?

Not all complications are equally important to patient safety. For instance, some complications have an extremely low rate of incidence (eg, pulmonary torsion), are not related to severe harm (eg, subcutaneous hematoma), or are not sensitive to prevention measures (eg, hypothyroidism after thyroidectomy). To ensure that the system is both operationally feasible and cost-effective in routine clinical practice, it is important to prioritize complications. We recommended prioritizing complications based on high incidence and serious prognosis and which are more likely to be preventable. We also suggested that it was important to consider local conditions. Table 1 shows an example from the MSCP Project, using empirical data and expert consensus to identify surgical complications with these 3 characteristics. Complications with low incidence can still be collected through the system. Increased experience and evidence may enable groups to identify underlining patterns for the occurrence of these complications and targeted intervention measures can then be formulated.

How Can We Translate Surgical Safety Information Into an Evidence-Based Stratifying Strategy?

An evidence-based stratifying strategy implies that patients with different risk factors (identified using data collected and consensus among the expert panel) will receive hierarchical and targeted interventions. The ability to make a preoperative determination of the overall risk for multiple major complications is a prerequisite for clinical decision making and securing surgical patients' safety, which is the ultimate goal of EPSIS [31-33]. Using the concept of *risk population* from epidemiology, we defined all surgical patients as the *risk population* for surgical complications at the stage before surgery [20]. In other words, every surgical patient could potentially develop any kind of complication and their risk of doing so is determined by factors that vary among the patients. Knowledge about patients' existing risk factors can be generated using a series of computational models to translate input data [34]. This means that major contributory factors to complications can be identified and intervention measures can then be developed for patients.

Oinas-Kukkonen argued that intervention should be "tailored to the potential needs, interests, personality, usage context, or other factors relevant to a user group" and that a system that offers personalized content or services has greater effectiveness and efficiency [35]. Patients can be managed through the system on the basis of their identified inherent risk. Specific interventions can be offered to patients with different levels of risk. This can, therefore, determine the appropriate amount of resources for each surgical patient, which is vital for optimizing patient flow. Examples of this concept include the well-known Physiological and Operative Severity Score for the enUmeration of Mortality and Morbidity scoring system and those recently developed by the National Surgical Quality Improvement Program in the United States [25,36]. With the growing tendency for medical professionals to become more specialized, using the risk population concept in the surgical safety field can also provide a macroscopic view across safety issues in multiple specialties. It, therefore, provides a more global picture for system-wide intervention planning.

How Can We Formulate Targeted Intervention Guidelines?

A system-integrated intervention guideline for decision making should map out which specific surgical safety intervention measures should be provided to patients with particular baseline risk and how the intervention should be carried out. The evidence from theory, the literature, and empirical research could be brought together and developed through standard procedures, such as the Delphi process [37]. Frontline staff and senior management, including both administrative and clinical leaders, should be involved in formulating the guidelines to get a buy-in from different professional groups (eg, surgeons, nurses, and anesthesiologists) to use the system to deliver interventions [14,38]. For instance, in MSCP, the interdisciplinary team of centralized researchers and clinicians reviewed the relevant research to identify interventions with the greatest benefit and the lowest barriers to use. Figure 2 shows how patients, complications, and interventions are provided in a stratified way in this system.

XSL•FO
RenderX

Table 1. Examples of prioritization of complications in the modern surgery and anesthesia safety management system construction and promotion project.

Complication	Decision
Surgical site infection	A[a]
Delayed healing or nonhealing incision	A
On ventilator ≥48 hours	A
Death or confirmed death	A
Hypoparathyroidism	B[b]
Coma after operation ≥24 hours	B
Respiratory failure	A
Electrolyte disturbance	B
Urinary tract infection	A
Pleural effusion	B
Acute renal failure	A
Skull defect	B
Cerebral edema	B
Hemorrhage requiring ≥4U RBC infusion 72 hours postoperatively	A
Iatrogenic pneumothorax	C[c]
Esophagus anastomotic fistula	C
Aphasia	C
Stress ulcer	C
Pneumocrania	C
Myasthenia crisis	C
Vocal cord paralysis	C
Pulmonary torsion	C
Secondary spinal canal stenosis	C
Fracture or loosening or dislocation of prosthesis	C
Internal or external fistula formation	C
Tracheal softening and collapse	C
Incisional hernia	C
Postoperative skin flap or subcutaneous effusion	C
Heterotopic ossification	C

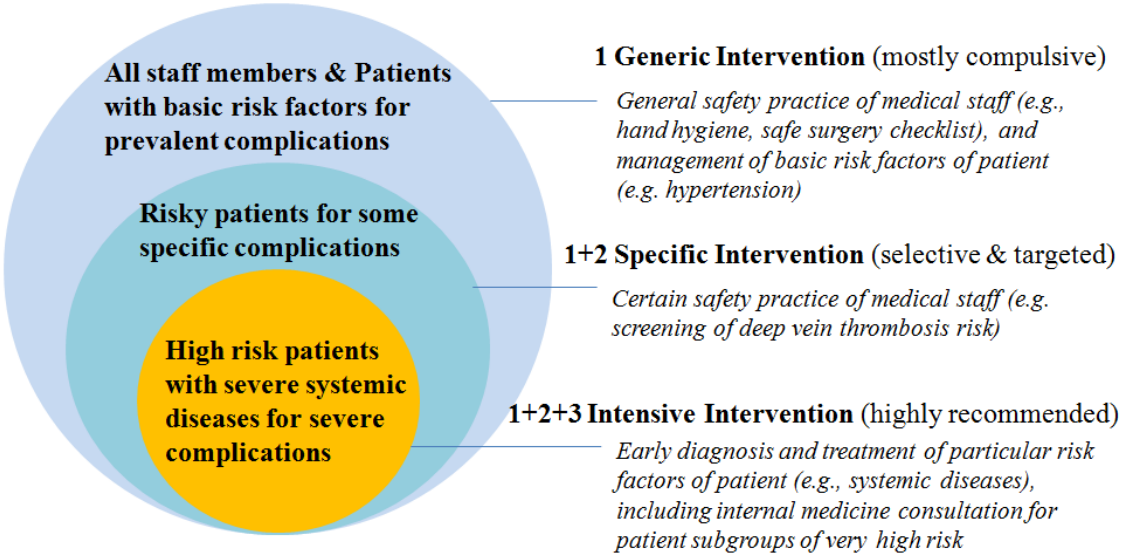
[a]High incidence (≥0.5%), severe harm, and sensitive to prevention measure.

[b]High incidence, but not considered severe, or preventable based on the literature and an expert consensus.

[c]Very low incidence, based on complications reported through the system.

XSL•FO
RenderX

Figure 2. Stratified intervention for patients at different risk level of complications.



Developing Functions and Applications

The data or applications should be presented in a visual form that users can understand. This makes them easy to use to support day-to-day operations and provide high-quality communication among hospital sectors for information- and knowledge-related functions [39]. Frontline medical staff and administrators at different levels should be extensively involved in the development process because adoption needs to be driven by clinicians where significant benefits can be articulated for them, the clinical and administrative teams, and the patients [40].

How Can We Integrate the System Into the Existing Clinical Workflow and Make the System Sustainable?

The objective of EPSIS is improving surgical safety by optimizing rather than subverting the current workflow. However, the information system may be frustrating for frontline clinicians and organizations if it does not fit with existing systems. This is particularly true if it causes longer completion times and workflow disruptions [24,41]. The system should, therefore, be embedded into existing clinical routines in line with the way frontline medical staff like to work. It should also take into account the interdependencies among the health care staff, cultural environment, and the infrastructural organization of the hospital. The challenge in achieving this goal is to identify the critical elements in the surgical workflow that influence patient safety management and the need to exchange safety information among medical staff at a minimum cost to existing workflows and thus provide information support for the next task. System use by different people in the clinical setting may be improved by visually representing the workflow of a complex clinical work environment and using user-system interaction analysis and complex design changes. A picture will also help in imposing the necessary workflow control and has been shown to be more successful in changing an unsafe plan or preventing the omission of essential interventions [42,43]. For instance, *soft* or *hard* stop functions can be incorporated in the system. Soft stops can alert clinicians if the intervention is not carried out in line with the guidelines, and hard stops alert clinicians and stop the process unless the intervention has been completed or an explanation has been provided to the central control point to override the interception. Soft and hard stops are helpful in promoting a buy-in but can result in variation in practice and poor compliance with safety goals and intervention measures [44]. An example of soft stops from MSCP is an interception function developed to ensure the timely delivery of preoperative patient safety interventions. The surgery submission is not approved if the preoperative intervention has not been completed for the patient.

How Can We Share Key Surgical Safety Information Smoothly and in a Timely Way?

Perioperative clinicians and staff have little opportunity to become familiar with surgical patients other than a quick determination of the required procedure. This lack of familiarity with and knowledge about patients could mean that perioperative team members (eg, operation room, postanesthesia care unit, and intensive care unit staff) might omit information that is important for surgical safety, such as allergies or antibiotic use. Integrated care and a high degree of interoperability should be highlighted to ensure high-quality communication of safety information among the various hospital sectors. This will allow quick and adequate responses about surgical patient safety issues [39]. Automatic reminders or alerts can facilitate the seamless transfer of vital information. They can also ensure that appropriate information is delivered to the surgeon at the right time and in a way that will ensure the surgeon receives and acts upon it. For instance, a reminder function was developed in MSCP to alert staff across the surgical ward and operation room when patients had high American Society of Anesthesiologists classifications.

How Can We Make the User Interface Acceptable to End Users?

The interface is one of the most significant parts of an information system and helps users to work efficiently, effectively, and satisfactorily [40,45]. The interface design should aim to eliminate complexity, emphasize key elements, and use special colors to mark important areas. One tactic is to

XSL•FO
RenderX

make the interface as similar as possible to the previous paper records, so that users do not need to search for the required fields. Only the field relevant to the current task should be displayed to the user, to increase efficiency. Feedback from end users should be sought and welcomed on an ongoing basis for the improvement of interface design and the iteration of development as a result.

How Can We Achieve a High Level of Interoperability With Existing Information Systems and What is the Benefit?

Hospitals are information technology–intensive workplaces, incorporating many kinds of information systems (eg, electronic medical records, laboratory information management system, and office automation systems). These form an interoperable digital health ecosystem. Integration of the EPSIS with existing information systems could, therefore, influence several layers of caregivers and provide more convenient workflow control. Caregivers and professionals will be able to send, receive, find, and use digital health and care information in an appropriate, secure, timely, and reliable way and with little additional effort. Data interface standards should be established in the data-sharing process for patient-level data.

In MSCP, we integrated our intervention system into the existing hospital information system, which, we believe, provided considerable advantages in terms of the extent, depth, and value-added use of the intervention system [46]. For instance, being able to change intervention information electronically makes it easy to continuously monitor and validate the intervention behavior of medical staff. Integration is also helpful in avoiding the inconvenience of switching between different systems and, therefore, decreasing staff resistance to the intervention system [47]. Other benefits include the relative low learning cost, faster adoption, and ease of logging-in. However, there are also some barriers to integration with other information systems, such as the level of investment needed, additional data leakage risk, higher maintenance cost, and administrative resistance.

System Implementation

There are two main challenges to managing the adoption, implementation, and sustainability of the system. The first is implementing the system organizationally, and the second is shaping the use of the system and related practices to achieve practical alignment with the intervention intention [48,49]. Drawing up an implementation plan and identifying and selecting appropriate methods or techniques that fit the context are considered fundamental to successful implementation [50].

What Contextual Factors Influence System Implementation?

Ideas, practices, organizational arrangements, roles, and status in the information system all reflect and are influenced by the wider sociocultural context in which they occur [51]. This is particularly true for the organizational setting within which an information system is implemented, because it forms an integral part of that system. Several frameworks are available to identify contextual factors that are likely to influence the implementation of a given intervention [52,53]. Their use allows attention to be directed toward the contextual factors that are likely to hinder implementation as well as identifying facilitators of success [54,55]. For instance, Meijden et al identified 6 dimensions affecting implementation: (1) system quality, (2) information quality, (3) usage, (4) user satisfaction, (5) individual impact, and (6) organizational impact [56].

How Should the Surgical Safety Culture and its Improvement Be Assessed?

Information system initiatives often fail because of mismatching between culture and the information system or a failure to understand culture and its influence on end user adoption of information systems [51]. Changing staff attitudes and views can help the staff understand the reasons for system changes and improve acceptance [57-59]. The use of tools, such as the Healthcare Research and Quality Hospital Survey on Patient Safety Culture, is recommended to inform perceptions about safety and related behavior, as well as to support the adoption of a safety intervention system [60-62]. However, culture change in a health care institution can be a long and intensive process requiring more cross-collaboration and greater user participation at all levels. Studies have suggested that organizational champions who can *shepherd* system implementation, influence cultural change, and act as a bridge with developers would be a valuable resource. Strategies such as printed educational materials, educational meetings, and educational outreach are also effective in changing attitudes and behaviors and increasing the use of a new information system [38,63].

Training Strategy

The implementation of a new information system in a clinical setting often ignores the influence of processes and routines of clinical practice. A lack of understanding of system capabilities can lead to workarounds, with the new system being used in unintended ways [64]. A training plan must be designed and completed before the initial implementation and should include intensive support during implementation. Moreover, providing training shows the organization's support in system implementation and development [65]. It is often effective to deliver training tailored to different users, for example, nurses, clinicians, and medical administrators. This should, however, include a holistic view of the entire system to strengthen the understanding of its function and goal. Hands-on practice and simulations may provide more benefit than only giving lectures [66]. It is also better to provide both compulsory and voluntary training elements [67]. Finally, training sessions are essential during or before system implementation, but ongoing training and development are also important [68]. Before the start of the MSCP project, training teams involving both clinicians and information technology engineers were established in each hospital. Unit-based education and training sessions were provided for nurses and other clinicians separately, because of the differences in workflow and the system operating interface. The standard operating procedure for efficient operation and compliance was translated into a course to facilitate the training process. Continuous training was also a part of MSCP, to reflect system updates and deployment of new function modules.

XSL•FO
RenderX

Evaluation and Continuous Improvement

Surgical safety improvement is an iterative procedure. Postimplementation evaluation, feedback, and performance gap assessment can deepen insight into how and why particular changes did or did not occur. This can further increase the benefits of using the system, prolong its sustainability, and support dynamic learning and improvement [69,70].

What Implementation Outcomes Should Be Used to Assess System Implementation?

When developing monitoring and evaluation plans for an EPSIS, new system implementation indicators (eg, usability) are recommended in addition to clinical outcomes (eg, complications, death, and length of stay). Different dimensions of outcomes and/or assessment methodologies have been proposed for evaluating the implementation of health care information systems [56,71-73]. For instance, Proctor et al distinguished among 3 distinct but interrelated types of outcomes for assessment in implementation studies: service (eg, safety), implementation (eg, fidelity), and client outcomes (eg, satisfaction) [73]. Hull et al outlined 8 implementation outcomes, defined as "the effects of deliberate and purposive actions to implement new treatments, practices, and services". The outcomes were acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration diffusion, and sustainability [74]. Correlating these implementation outcomes with implementation success and failure could strengthen the understanding of the mechanisms of the effect. This might further translate into evidence-based interventions to improve surgical safety.

Monitoring, Auditing, and Formulating a Feedback Channel for Performance Improvement

A surgical safety information system needs monitoring and continuous auditing to ensure that it adheres to intervention guidelines [75]. Messaging functions could be used to enable real-time recording and transmission of any problems. In MSCP, the completion status of the surgical safety checklist for the operation room is recorded at different stages (sign in, time out, and sign out), and any violation at any stage is recorded and submitted.

Establishing a mechanism to provide feedback on the results of intervention protocol variations (eg, compliance and completeness) and patient outcomes to frontline medical staff and managers is essential for fostering a culture of surgical patient safety in hospitals [76,77]. These measures can encourage improvements in identifying and sharing information about patient safety incidents. They can also help the staff to identify problems during the implementation process, ultimately supporting continuous learning as well as increasing engagement of medical staff. For instance, case-enhanced learning can use real cases to allow staff to identify problems and solutions. This, therefore, provides learning resources about complications and clinical behaviors to support safety improvement. A redesign-action-feedback closed cycle can be formed among knowledge generation, function and application development, system implementation, and evaluation of effectiveness. This has no defined beginning and end point, and the cycle should not be interpreted as starting with prevention and ending with action.

In the MSCP project, a *graded confirmed awareness* subsystem was created. This gradually summarized patient outcomes and the implementation status of the system to improve awareness and confirmation in a specific sequence. This moved from patient to doctor and nurse in charge, to surgeon (daily), to department head (weekly), to medical affairs department (monthly), to hospital dean (quarterly), and then to permanent database. This system established dual duties (for subordinates and superiors) and dual supervision mechanisms (individual and external) for each person. Frontline staff and managers become equal elements in the chain securing patient safety. More importantly, this type of feedback channel could also help create a culture of patient safety and increase the motivation of hospital staff to deliver interventions through the system. In the MSCP, higher implementation rates were observed for all the postoperative prevention packages following the introduction of this subsystem (Table 2).

Table 2. Effects of implementing a Graded Confirmed Awareness System on surgeons in 4 hospitals in the modern surgery and anesthesia safety management system construction and promotion project.

Prevention category	Number with risk factors	Preoperative prevention measures delivered	Postoperative prevention measures delivered
Before introduction of the graded confirmed awareness system, n (%)			
Generic intervention	11,472 (27.0)	10,671 (93.0)	5830 (50.8)
Specific intervention	11,641 (31.7)	11,453 (98.4)	6905 (59.3)
Intensive intervention	4588 (23.8)	3245 (70.7)	2064 (44.5)
After introduction of the graded confirmed awareness system, n (%)			
Generic intervention	8947 (23.6)	7528 (84.1)	5535 (61.9)
Specific intervention	8532 (23.2)	8270 (96.9)	6037 (70.8)
Intensive intervention	3514 (20.1)	2241 (63.8)	1709 (48.6)

XSL•FO
RenderX

Conclusions

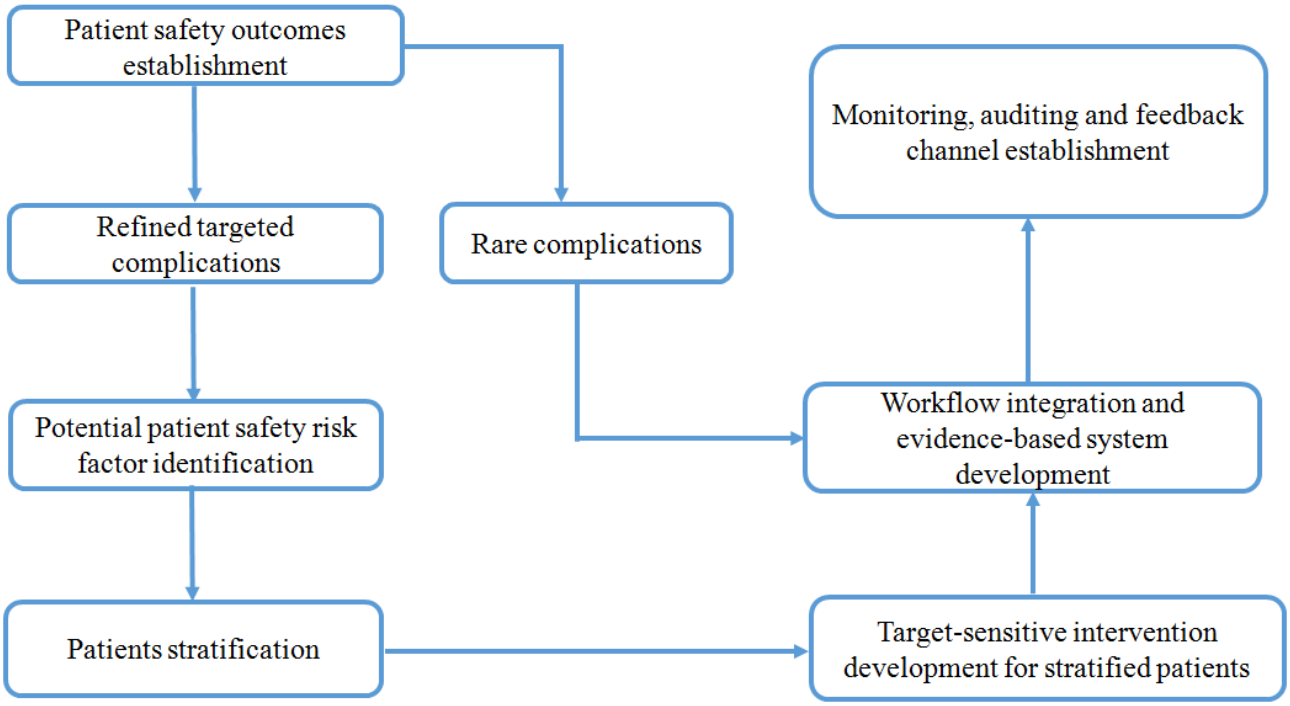
This study has outlined and discussed several critical issues in the process of developing an EPSIS. The design rationale of the system is summarized in Figure 3. The EPSIS is characterized by several distinct features, including the following:

1. Optimizing patient flow management and avoiding overload of medical staff by delivering tailored interventions to patients in a target-sensitive and stratified way based on inherent risk.
2. Incorporating the intervention into routine clinical activities and formulating a surgical safety information circle that ensures that critical safety and management information transfers smoothly between sectors and medical staff, in particular, senior managers, who play a pivotal role in this information circle.
3. Integrating the intervention information system with the existing hospital information system to construct an interoperable surgical safety improvement ecosystem, incorporating outcome reporting, intervention, monitoring, and feedback and finally developing a surgical safety learning system encouraging a patient safety culture in the hospital.

The issues we have discussed in this paper might serve as a reference for projects aiming to build an interactive mechanism to combine safety systems supporting wide and continuous improvement in surgical safety.

There are a number of limitations and challenges in constructing the EPSIS. The EPSIS in MSCP was mainly for patients admitted for elective surgery. Future expansion, for example, to emergency surgery and intensive care units, will present other challenges that have not been considered in this paper. It is difficult to establish evidence-based interventions for complications with low incidence in EPSIS, which is focused on complications. However, EPSIS is both a management system and learning platform. As more information is gathered, this problem will gradually be eliminated.

Figure 3. Design pattern rationale graph of an evidence-based surgical safety information system.



Acknowledgments

This study was funded by the National Health and Family Planning Commission of China (grant number: 201402017). The authors thank all the medical staff members and experts for their contributions of time, suggestions, and encouragement in this study.

Authors' Contributions

All authors listed in our manuscript have contributed substantially to this study. YZ, as the corresponding author, was responsible for constructing the idea and overarching research goals and aims. XY, WH, and JJ were responsible for designing and coordinating the research, supervising the course, and drafting the manuscript. YW, SX, SW, and HS were responsible for reviewing the paper for its intellectual content before submission. ZW was responsible for the literature review and also took responsibility for reviewing the paper for spelling and grammar.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Weiser TG, Makary MA, Haynes AB, Dziekan G, Berry WR, Gawande AA, Safe Surgery Saves Lives Measurement and Study Groups. Standardised metrics for global surgical surveillance. Lancet 2009 Sep 26;374(9695):1113-1117. [doi: 10.1016/S0140-6736(09)61161-2] [Medline: 19782877]
2. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, Safe Surgery Saves Lives Study Group. A surgical safety checklist to reduce morbidity and mortality in a global population. N Engl J Med 2009 Jan 29;360(5):491-499. [doi: 10.1056/NEJMsa0810119] [Medline: 19144931]
3. Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, et al. Public reporting and pay for performance in hospital quality improvement. N Engl J Med 2007 Feb 01;356(5):486-496. [doi: 10.1056/NEJMsa064964] [Medline: 17259444]
4. Agrawal A. Medication errors: prevention using information technology systems. Br J Clin Pharmacol 2009 Jun;67(6):681-686 [FREE Full text] [doi: 10.1111/j.1365-2125.2009.03427.x] [Medline: 19594538]
5. Batra R, Wolbach-Lowes J, Swindells S, Scarsi KK, Podany AT, Sayles H, et al. Impact of an electronic medical record on the incidence of antiretroviral prescription errors and HIV pharmacist reconciliation on error correction among hospitalized HIV-infected patients. Antivir Ther 2015;20(5):555-559. [doi: 10.3851/IMP2930] [Medline: 25560773]
6. Stultz JS, Nahata MC. Preventability of Voluntarily Reported or Trigger Tool-Identified Medication Errors in a Pediatric Institution by Information Technology: A Retrospective Cohort Study. Drug Saf 2015 Jul;38(7):661-670. [doi: 10.1007/s40264-015-0303-y] [Medline: 26013909]
7. Padhy RP, Patra MR, Satapathy SC. Design and implementation of a cloud based rural healthcare information system model. UNIASCIT 2012;2(1):149-157 [FREE Full text]
8. Carter AW, Mossialos E, Darzi A. A national incident reporting and learning system in England and Wales, but at what cost? Expert Rev Pharmacoecon Outcomes Res 2015 Jun;15(3):365-368. [doi: 10.1586/14737167.2015.1031114] [Medline: 25834922]
9. Institute of Medicine (US) Committee on Quality of Health Care in America, Kohn LT, Corrigan JM, Donaldson MS. To err is human: building a safer health system. Washington (DC): National Academies Press (US) 2000:1-287. [doi: 10.17226/9728] [Medline: 25077248]
10. Meguid RA, Bronsert MR, Juarez-Colunga E, Hammermeister KE, Henderson WG. Surgical risk preoperative assessment system (SURPAS): II. Parsimonious risk models for postoperative adverse outcomes addressing need for laboratory variables and surgeon specialty-specific models. Ann Surg 2016 Jul;264(1):10-22. [doi: 10.1097/SLA.0000000000001677] [Medline: 26945154]
11. Marang-van de Mheen PJ, Kievit J. Automated registration of adverse events in surgical patients in the Netherlands: the current status. Ned Tijdschr Geneeskd 2003 Jun 28;147(26):1273-1277. [Medline: 12861669]
12. Ozrazgat-Baslanti T, Blanc P, Thottakkara P, Ruppert M, Rashidi P, Momcilovic P, et al. Preoperative assessment of the risk for multiple complications after surgery. Surgery 2016 Aug;160(2):463-472 [FREE Full text] [doi: 10.1016/j.surg.2016.04.013] [Medline: 27238354]
13. Tingle J. Patient safety at the crossroads. Br J Nurs 2017 Nov 09;26(20):1138-1139. [doi: 10.12968/bjon.2017.26.20.1138] [Medline: 29125339]
14. Urbach DR, Govindarajan A, Saskin R, Wilton AS, Baxter NN. Introduction of surgical safety checklists in Ontario, Canada. N Engl J Med 2014 Mar 13;370(11):1029-1038. [doi: 10.1056/NEJMsa1308261] [Medline: 24620866]
15. Baek H, Cho M, Kim S, Hwang H, Song M, Yoo S. Analysis of length of hospital stay using electronic health records: a statistical and data mining approach. PLoS One 2018 Apr 13;13(4):e0195901 [FREE Full text] [doi: 10.1371/journal.pone.0195901] [Medline: 29652932]
16. Pollak VE, Lorch JA. Effect of electronic patient record use on mortality in End Stage Renal Disease, a model chronic disease: retrospective analysis of 9 years of prospectively collected data. BMC Med Inform Decis Mak 2007 Nov 28;7:38. [doi: 10.1186/1472-6947-7-38] [Medline: 18045495]
17. Abraham C, Junglas IA. From cacophony to harmony: a case study about the IS implementation process as an opportunity for organizational transformation at Sentara Healthcare. J Strategic Inf Syst 2011 Jun;20(2):177-197. [doi: 10.1016/j.jsis.2011.03.005]
18. Schimpff SC. Improving operating room and perioperative safety: background and specific recommendations. Surg Innov 2007 Jun;14(2):127-135. [doi: 10.1177/1553350607301746] [Medline: 17558019]
19. Ulrich B, Kear T. Patient safety and patient safety culture: foundations of excellent health care delivery. Nephrol Nurs J 2014 Sep;41(5):447-456. [Medline: 26295088]
20. Yu X, Jiang J, Liu C, Shen K, Wang Z, Han W, et al. Protocol for a multicentre, multistage, prospective study in China using system-based approaches for consistent improvement in surgical safety. BMJ Open 2017 Jun 15;7(6):e015147 [FREE Full text] [doi: 10.1136/bmjopen-2016-015147] [Medline: 28619774]

XSL•FO
RenderX

21. Bärnighausen T, Bloom DE, Humair S. Health systems and HIV treatment in sub-Saharan Africa: matching intervention and programme evaluation strategies. Sex Transm Infect 2012 Mar;88(2):e2. [doi: 10.1136/sextrans-2011-050303] [Medline: 22345032]
22. Lilford RJ, Chilton PJ, Hemming K, Girling AJ, Taylor CA, Barach P. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. Br Med J 2010 Aug 27;341:c4413. [doi: 10.1136/bmj.c4413] [Medline: 20802000]
23. Archer S, Hull L, Soukup T, Mayer E, Athanasiou T, Sevdalis N, et al. Development of a theoretical framework of factors affecting patient safety incident reporting: a theoretical review of the literature. BMJ Open 2017 Dec 27;7(12):e017155 [FREE Full text] [doi: 10.1136/bmjopen-2017-017155] [Medline: 29284714]
24. Sheikh A, Cornford T, Barber N, Avery A, Takian A, Lichtner V, et al. Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from prospective national evaluation in. Br Med J 2011 Oct 17;343:d6054 [FREE Full text] [doi: 10.1136/bmj.d6054] [Medline: 22006942]
25. Khuri SF, Daley J, Henderson W, Hur K, Demakis J, Aust JB, et al. The Department of Veterans Affairs' NSQIP: the first national, validated, outcome-based, risk-adjusted, and peer-controlled program for the measurement and enhancement of the quality of surgical care. National VA Surgical Quality Improvement Program. Ann Surg 1998 Oct;228(4):491-507. [doi: 10.1097/00000658-199810000-00006] [Medline: 9790339]
26. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999 Jul;16(1):9-13. [doi: 10.1016/s1010-7940(99)00134-7] [Medline: 10456395]
27. Tsang C, Palmer W, Bottle A, Majeed A, Aylin P. A review of patient safety measures based on routinely collected hospital data. Am J Med Qual 2012;27(2):154-169. [doi: 10.1177/1062860611414697] [Medline: 21896785]
28. Clavien PA, Sanabria JR, Strasberg SM. Proposed classification of complications of surgery with examples of utility in cholecystectomy. Surgery 1992 May;111(5):518-526. [Medline: 1598671]
29. Martin RC, Brennan MF, Jaques DP. Quality of complication reporting in the surgical literature. Ann Surg 2002 Jun;235(6):803-813. [doi: 10.1097/00000658-200206000-00007] [Medline: 12035036]
30. Reiner B. Strategies for medical data extraction and presentation part 2: creating a customizable context and user-specific patient reference database. J Digit Imaging 2015 Jun;28(3):249-255 [FREE Full text] [doi: 10.1007/s10278-015-9794-4] [Medline: 25833767]
31. Schoenfeld AJ, Ochoa LM, Bader JO, Belmont PJ. Risk factors for immediate postoperative complications and mortality following spine surgery: a study of 3475 patients from the National Surgical Quality Improvement Program. J Bone Joint Surg Am 2011 Sep 07;93(17):1577-1582. [doi: 10.2106/JBJS.J.01048] [Medline: 21915571]
32. Saxton A, Velanovich V. Preoperative frailty and quality of life as predictors of postoperative complications. Ann Surg 2011 Jun;253(6):1223-1229. [doi: 10.1097/SLA.0b013e318214bce7] [Medline: 21412145]
33. Vogel TR, Dombrovskiy VY, Carson JL, Graham AM, Lowry SF. Postoperative Sepsis in the United States. Ann Surg 2010 Dec;252(6):1065-1071. [doi: 10.1097/SLA.0b013e3181dcf36e] [Medline: 20571363]
34. Dodig-Crnkovic G. Knowledge generation as natural computation. J Syst Cybern Inf 2008;6(2):12-16.
35. Wahle F, Bollhalder L, Kowatsch T, Fleisch E. Toward the design of evidence-based mental health information systems for people with depression: a systematic literature review and meta-analysis. J Med Internet Res 2017 May 31;19(5):e191 [FREE Full text] [doi: 10.2196/jmir.7381] [Medline: 28566267]
36. Brooks MJ, Sutton R, Sarin S. Comparison of surgical risk score, POSSUM and p-POSSUM in higher-risk surgical patients. Br J Surg 2005 Oct;92(10):1288-1292. [doi: 10.1002/bjs.5058] [Medline: 15981213]
37. Jones J, Hunter D. Consensus methods for medical and health services research. Br Med J 1995 Aug 05;311(7001):376-380 [FREE Full text] [doi: 10.1136/bmj.311.7001.376] [Medline: 7640549]
38. Bergs J, Lambrechts F, Simons P, Vlayen A, Marneffe W, Hellings J, et al. Barriers and facilitators related to the implementation of surgical safety checklists: a systematic review of the qualitative evidence. BMJ Qual Saf 2015 Dec;24(12):776-786. [doi: 10.1136/bmjqs-2015-004021] [Medline: 26199428]
39. Winter AF, Ammenwerth E, Bott OJ, Brigl B, Buchauer A, Gräber S, et al. Strategic information management plans: the basis for systematic information management in hospitals. Int J Med Inform 2001 Dec;64(2-3):99-109. [Medline: 11734379]
40. Scholl J, Syed-Abdul S, Ahmed LA. A case study of an EMR system at a large hospital in India: challenges and strategies for successful adoption. J Biomed Inform 2011 Dec;44(6):958-967 [FREE Full text] [doi: 10.1016/j.jbi.2011.07.008] [Medline: 21846508]
41. Meyer KE, Altamore R, Chapko M, Miner M, McGann M, Hill E, et al. The need for a thoughtful deployment strategy: evaluating clinicians' perceptions of critical deployment issues. Stud Health Technol Inform 1998;52 Pt 2:854-858. [Medline: 10384582]
42. Strom BL, Schinnar R, Aberra F, Bilker W, Hennessy S, Leonard CE, et al. Unintended effects of a computerized physician order entry nearly hard-stop alert to prevent a drug interaction: a randomized controlled trial. Arch Intern Med 2010 Sep 27;170(17):1578-1583. [doi: 10.1001/archinternmed.2010.324] [Medline: 20876410]

XSL•FO
RenderX

43. Eslami S, de Keizer NF, Abu-Hanna A. The impact of computerized physician medication order entry in hospitalized patients—a systematic review. Int J Med Inform 2008 Jun;77(6):365-376. [doi: 10.1016/j.ijmedinf.2007.10.001] [Medline: 18023611]
44. Feldman SS, Buchalter S, Hayes LW. Health information technology in healthcare quality and patient safety: literature review. JMIR Med Inform 2018 Jun 04;6(2):e10264 [FREE Full text] [doi: 10.2196/10264] [Medline: 29866642]
45. Jitnupong B, Jirachiefpattana W. Information system user interface design in software services organization: a small-clan case study. MATEC Web Conf 2018 Apr 23;164:01006. [doi: 10.1051/matecconf/201816401006]
46. Desroches CM, Agarwal R, Angst CM, Fischer MA. Differences between integrated and stand-alone E-prescribing systems have implications for future use. Health Aff (Millwood) 2010 Dec;29(12):2268-2277. [doi: 10.1377/hlthaff.2010.0557] [Medline: 21134929]
47. McDonald CJ. The barriers to electronic medical record systems and how to overcome them. J Am Med Inform Assoc 1997;4(3):213-221 [FREE Full text] [Medline: 9147340]
48. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. Med J Aust 2004 Mar 15;180(6 Suppl):S57-S60. [Medline: 15012583]
49. Arvidsson V, Holmström J, Lyytinen K. Information systems use as strategy practice: a multi-dimensional view of strategic information system implementation and use. J Strategic Inf Sys 2014 Mar;23(1):45-61. [doi: 10.1016/j.jsis.2014.01.004]
50. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013 Dec 01;8:139 [FREE Full text] [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
51. Jackson S. Organizational culture and information systems adoption: a three-perspective approach. Inf Organ 2011 Apr;21(2):57-83. [doi: 10.1016/j.infoandorg.2011.03.003]
52. Keith RE, Crosson JC, O'Malley AS, Cromp D, Taylor EF. Using the Consolidated Framework for Implementation Research (CFIR) to produce actionable findings: a rapid-cycle evaluation approach to improving implementation. Implement Sci 2017 Feb 10;12(1):15 [FREE Full text] [doi: 10.1186/s13012-017-0550-7] [Medline: 28187747]
53. Ward MM, Baloh J, Zhu X, Stewart GL. Promoting action on research implementation in health services framework applied to TeamSTEPPS implementation in small rural hospitals. Health Care Manage Rev 2017 Jan;42(1):2-13. [doi: 10.1097/HMR.0000000000000086] [Medline: 26415078]
54. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
55. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
56. Van Der Meijden MJ, Tange HJ, Troost J, Hasman A. Determinants of success of inpatient clinical information systems: a literature review. J Am Med Inform Assoc 2003;10(3):235-243 [FREE Full text] [doi: 10.1197/jamia.M1094] [Medline: 12626373]
57. Borchard A, Schwappach DL, Barbir A, Bezzola P. A systematic review of the effectiveness, compliance, and critical factors for implementation of safety checklists in surgery. Ann Surg 2012 Dec;256(6):925-933. [doi: 10.1097/SLA.0b013e3182682f27] [Medline: 22968074]
58. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, Safe Surgery Saves Lives Study Group. Changes in safety attitude and relationship to decreased postoperative morbidity and mortality following implementation of a checklist-based surgical safety intervention. BMJ Qual Saf 2011 Jan;20(1):102-107. [doi: 10.1136/bmjqs.2009.040022] [Medline: 21228082]
59. Parand A, Burnett S, Benn J, Iskander S, Pinto A, Vincent C. Medical engagement in organisation-wide safety and quality-improvement programmes: experience in the UK Safer Patients Initiative. Qual Saf Health Care 2010 Oct;19(5):e44. [doi: 10.1136/qshc.2009.036368] [Medline: 20538629]
60. de Wet C, Johnson P, Mash R, McConnachie A, Bowie P. Measuring perceptions of safety climate in primary care: a cross-sectional study. J Eval Clin Pract 2012 Feb;18(1):135-142. [doi: 10.1111/j.1365-2753.2010.01537.x] [Medline: 20860593]
61. Flin R, Burns C, Mearns K, Yule S, Robertson EM. Measuring safety climate in health care. Qual Saf Health Care 2006 Apr;15(2):109-115 [FREE Full text] [doi: 10.1136/qshc.2005.014761] [Medline: 16585110]
62. Nieva VF, Sorra J. Safety culture assessment: a tool for improving patient safety in healthcare organizations. Qual Saf Health Care 2003 Dec;12 Suppl 2:ii17-ii23 [FREE Full text] [doi: 10.1136/qhc.12.suppl_2.ii17] [Medline: 14645891]
63. Russ SJ, Sevdalis N, Moorthy K, Mayer EK, Rout S, Caris J, et al. A qualitative evaluation of the barriers and facilitators toward implementation of the WHO surgical safety checklist across hospitals in England: lessons from the. Ann Surg 2015 Jan;261(1):81-91. [doi: 10.1097/SLA.0000000000000793] [Medline: 25072435]
64. Cresswell KM, Bates DW, Sheikh A. Ten key considerations for the successful implementation and adoption of large-scale health information technology. J Am Med Inform Assoc 2013 Jun;20(e1):e9-e13 [FREE Full text] [doi: 10.1136/amiajnl-2013-001684] [Medline: 23599226]

XSL•FO
RenderX

65. Anderson LK, Stafford CJ. The "big bang" implementation: not for the faint of heart. Comput Nurs 2002;20(1):14-20; quiz 20. [Medline: 11838384]
66. Milano CE, Hardman JA, Plesiu A, Rdesinski RE, Biagioli FE. Simulated electronic health record (Sim-EHR) curriculum: teaching EHR skills and use of the EHR for disease management and prevention. Acad Med 2014 Mar;89(3):399-403. [doi: 10.1097/ACM.0000000000000149] [Medline: 24448035]
67. Gegenfurtner A, Könings KD, Kosmajac N, Gebhardt M. Voluntary or mandatory training participation as a moderator in the relationship between goal orientations and transfer of training. Int J Train Dev 2016 Oct 03;20(4):290-301. [doi: 10.1111/ijtd.12089] [Medline: 24448035]
68. Ajami S, Mohammadi-Bertiani Z. Training and its Impact on Hospital Information System (HIS) Success. J Inform Tech Softw Eng 2012 Jan;2:112. [doi: 10.4172/2165-7866.1000112]
69. Berwick DM. Measuring surgical outcomes for improvement: was Codman wrong? J Am Med Assoc 2015 Feb 03;313(5):469-470. [doi: 10.1001/jama.2015.4] [Medline: 25647202]
70. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012 Jun 13(6):CD000259. [doi: 10.1002/14651858.CD000259.pub3] [Medline: 22696318]
71. Nelson L, Taylor F, Adams M, Parker DE. Improving pain management for hip fractured elderly. Orthop Nurs 1990;9(3):79-83. [doi: 10.1097/00006416-199005000-00012] [Medline: 2377382]
72. Stoop AP, Berg M. Integrating quantitative and qualitative methods in patient care information system evaluation: guidance for the organizational decision maker. Methods Inf Med 2003;42(4):458-462. [Medline: 14534650]
73. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
74. Hull L, Athanasiou T, Russ S. Implementation science: a neglected opportunity to accelerate improvements in the safety and quality of surgical care. Ann Surg 2017 Jun;265(6):1104-1112. [doi: 10.1097/SLA.0000000000002013] [Medline: 27735828]
75. Mutale W, Chintu N, Amoroso C, Awoonor-Williams K, Phillips J, Baynes C, Population Health Implementation and Training – Africa Health Initiative Data Collaborative. Improving health information systems for decision making across five sub-Saharan African countries: Implementation strategies from the African Health Initiative. BMC Health Serv Res 2013;13 Suppl 2:S9 [FREE Full text] [doi: 10.1186/1472-6963-13-S2-S9] [Medline: 23819699]
76. Weaver SJ, Lubomksi LH, Wilson RF, Pfoh ER, Martinez KA, Dy SM. Promoting a culture of safety as a patient safety strategy: a systematic review. Ann Intern Med 2013 Mar 05;158(5 Pt 2):369-374 [FREE Full text] [doi: 10.7326/0003-4819-158-5-201303051-00002] [Medline: 23460092]
77. Latif A, Holzmueller CG, Pronovost PJ. Evaluating safety initiatives in healthcare. Curr Anesthesiol Rep 2014 Jun;4(2):100-106 [FREE Full text] [doi: 10.1007/s40140-014-0059-4] [Medline: 24976794]

Abbreviations

EPSIS: evidence-based patient safety information system
MSCP: modern surgery and anesthesia safety management system construction and promotion

Edited by G Eysenbach; submitted 06.02.19; peer-reviewed by F Graur, J Farzi; comments to author 28.03.19; revised version received 18.05.19; accepted 18.05.19; published 24.06.19.

Please cite as:
Yu X, Han W, Jiang J, Wang Y, Xin S, Wu S, Sun H, Wang Z, Zhao Y
Key Issues in the Development of an Evidence-Based Stratified Surgical Patient Safety Improvement Information System: Experience From a Multicenter Surgical Safety Program
J Med Internet Res 2019;21(6):e13576
URL: http://www.jmir.org/2019/6/e13576/
doi:10.2196/13576
PMID:31237241

©Xiaochu Yu, Wei Han, Jingmei Jiang, Yipeng Wang, Shijie Xin, Shizheng Wu, Hong Sun, Zixing Wang, Yupei Zhao. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 24.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of

XSL•FO
RenderX

Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effectiveness of an App-Based Nurse-Moderated Program for New Mothers With Depression and Parenting Problems (eMums Plus): Pragmatic Randomized Controlled Trial

Alyssa Sawyer[1,2,3], PhD (ClinPsy); Amy Kaim[2,3], BHlthSci (Hons); Huynh-Nhu Le[4], PhD; Denise McDonald[5], BSc (Community Health Nursing); Murthy Mittinty[1], PhD; John Lynch[1,6], PhD; Michael Sawyer[2,3], PhD

[1]School of Public Health, University of Adelaide, Adelaide, Australia
[2]School of Medicine, University of Adelaide, Adelaide, Australia
[3]Research and Evaluation Unit, Women's and Children's Health Network, Adelaide, Australia
[4]Department of Psychology, George Washington University, Washington DC, DC, United States
[5]Child and Family Health Service, Women's and Children's Health Network, Adelaide, Australia
[6]Population Health Sciences, University of Bristol, Bristol, United Kingdom

Corresponding Author:
Alyssa Sawyer, PhD (ClinPsy)
School of Public Health
University of Adelaide
Adelaide Health & Medical Sciences Building
57 North Terrace, Mail Drop DX 650 550
Adelaide, 5005
Australia
Phone: 61 8 81617207
Fax: 61 8 81616906
Email: alyssa.sawyer@adelaide.edu.au

Abstract

Background: Postnatal depression and caregiving difficulties adversely affect mothers, infants, and later childhood development. In many countries, resources to help mothers and infants are limited. Online group–based nurse-led interventions have the potential to help address this problem by providing large numbers of mothers with access to professional and peer support during the postnatal period.

Objective: This study tested the effectiveness of a 4-month online group–based nurse-led intervention delivered when infants were aged 2 to 6 months as compared with standard care outcomes.

Methods: The study was a block randomized control trial. Mothers were recruited at the time they were contacted for the postnatal health check offered to all mothers in South Australia. Those who agreed to participate were randomly assigned to the intervention or standard care. The overall response rate was 63.3% (133/210). Primary outcomes were the level of maternal depressive symptoms assessed with the Edinburgh Postnatal Depression Scale (EPDS) and quality of maternal caregiving assessed using the Parenting Stress Index (PSI; competence and attachment subscales), the Parenting Sense of Competence Scale (PSCS), and the Nursing Child Assessment Satellite Training Scale. Assessments were completed at baseline (*mean child age* 4.9 weeks [*SD* 1.4]) and again when infants were aged 8 and 12 months.

Results: Outcomes were evaluated using linear generalized estimating equations adjusting for postrandomization group differences in demographic characteristics and the outcome score at baseline. There were no significant differences in the intervention and standard care groups in scores on the PSI competence subscale (*P*=.69) nor in the PSCS (*P*=.11). Although the group by time interaction suggested there were differences over time between the EPDS and PSI attachment subscale scores in the intervention and standard care groups (*P*=.001 and *P*=.04, respectively), these arose largely because the intervention group had stable scores over time whereas the standard care group showed some improvements between baseline and 12 months. Mothers engaged well with the intervention with at least 60% (43/72) of mothers logging-in once per week during the first 11 weeks of the intervention. The majority of mothers also rated the intervention as helpful and user-friendly.

XSL•FO
RenderX

Conclusions: Mothers reported that the intervention was helpful, and the app was described as easy to use. As such, it appears that support for mothers during the postnatal period, provided using mobile phone technology, has the potential to be an important addition to existing services. Possible explanations for the lack of differences in outcomes for the 2 groups in this study are the failure of many mothers to use key components of the intervention and residual differences between the intervention and standard care groups post randomization.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12616001732471; http://www.ANZCTR.org.au/ACTRN12616001732471.aspx (archived on WebCite as http://www.webcitation.org/77zo30GDw)

(J Med Internet Res 2019;21(6):e13689) doi:10.2196/13689

KEYWORDS

mobile phone; infant; mother-child relations; postnatal depression; randomized controlled trial

Introduction

Postnatal depression causes significant distress for mothers and is associated with a range of adverse outcomes for children. There is good evidence that caregiving difficulties associated with depressive symptoms play a key role in mediating the association between maternal depression and child outcomes [1-6]. For example, postnatal depression has the potential to negatively affect breastfeeding, infant sleep routines, and attendance at health checks, all of which can adversely affect children's later growth and development [4].

More recently it has been recognized that subthreshold postnatal depressive symptoms can also affect maternal functioning and infant development [6-8]. This has led many countries, including Australia, to initiate universal screening programs for early identification of mothers with depressive symptoms. However, a major challenge for screening programs is the paucity of clinical services available to support mothers with comorbid depressive symptoms and parenting problems and the difficulty of engaging busy and/or isolated new mothers with treatment programs [9].

Greater use of the internet to deliver services to mothers during the postnatal period offers a potential solution to these problems. Internet use by women of childbearing age in Australia and other countries is now ubiquitous with new mothers making extensive use of the internet to obtain child-raising information and social support [10-12]. This has encouraged the development of a vast array of websites and *mobile phone apps* by commercial, professional, and government organizations. However, as noted by Plantin and Daneback [10], health-related information on the internet can be misleading and occasionally, *utterly wrong* [10,13]. In addition, there is a marked absence of evaluations assessing whether online information and support can improve maternal and child outcomes.

The intervention evaluated in this study was developed as a response to the unresolved issue about how services can effectively address the high prevalence of mild-to-moderate depression and associated parenting problems experienced by many women during the postnatal period [14-16]. The study also responds to the National Institute for Health and Care Excellence guidelines calling for randomized controlled trials (RCTs) designed to evaluate the effectiveness of interventions to help mothers with depressive symptoms and parenting difficulties [1,6,7]. The latter is important as the subthreshold levels of postnatal depressive symptoms in the early postnatal period are a risk factor for the development of postnatal depression and also have the potential to interfere with optimal mother-infant development [6,17,18]. Importantly, an RCT is pragmatic and delivered within a state-wide child and family health service system. This approach was adopted to enhance the likelihood that translation into new service models would occur if the intervention was effective, as compared with efficacy trials conducted by researchers external to real-world service delivery systems.

The intervention extends earlier research by our group which evaluated an online group–based nurse-led intervention designed to help mothers in the general population better manage common parenting problems experienced by mothers during the postnatal period [19-21]. The enhanced intervention reported here used a 4-month online group–based nurse-led approach to provide help with both postnatal depressive symptoms and parenting problems for mothers who were identified as having comorbid mild-to-moderate levels of depressive symptoms and problems caring for their infants during the first 4 weeks after the birth of their infant.

The evaluation compared maternal and infant outcomes for those who received the new internet-based intervention versus outcomes for those who received standard postnatal home- and clinic-based support from a community nurse, routinely offered to all mothers in South Australia (Australian New Zealand Clinical Trials Registry ACTRN12616001732471). The primary outcomes were (1) the level of maternal depressive symptoms, (2) the quality of maternal caregiving including the level of parenting self-competence, and (3) the quality of the mother-infant relationship, assessed when infants were aged 8 and 12 months [22].

Methods

Participants, Recruitment, and Randomization

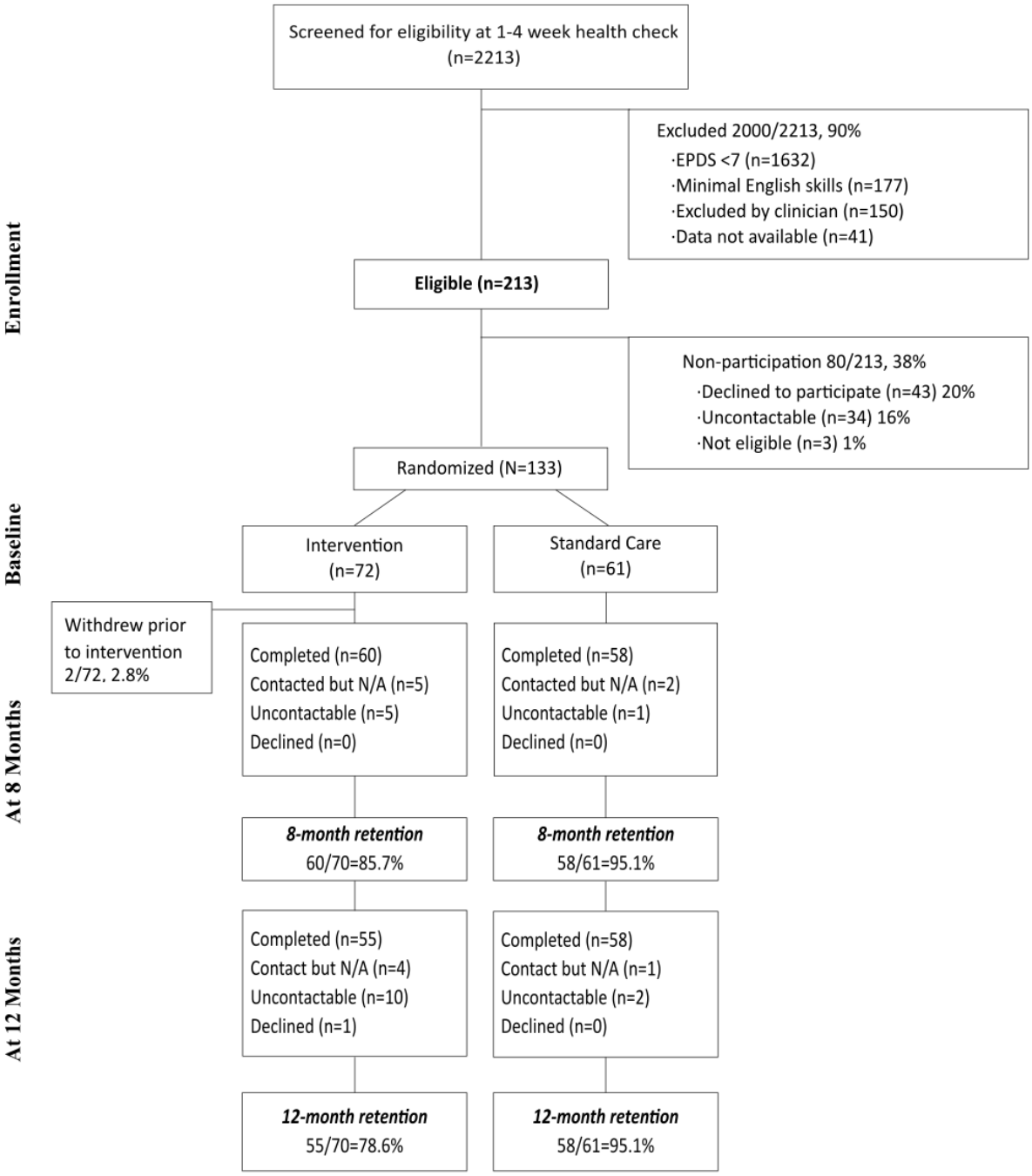
Participants were new mothers referred by their birthing hospital for their initial postnatal health check by nurses based at one of 13 Child and Family Health Service (CaFHS) community clinics located in Adelaide or 1 clinic located in a large regional center in South Australia. From March to June 2017, when nurses visited mothers for maternal postnatal health checks (undertaken when infants are aged 1 to 4 weeks), they asked mothers if they would give permission for the research team to contact them by telephone if the mother was eligible to participate in the

XSL•FO
RenderX

study. Eligibility criteria were (1) Edinburgh Postnatal Depression Scale (EPDS) score ≥7; (2) at least 1 self-reported parenting problem; (3) literacy in English; and (4) access to a smartphone.

In total, 2213 mothers agreed to be contacted by telephone and subsequently completed a nurse-provided questionnaire comprised of a 4-item parenting problem scale and the EPDS [23]. Among these mothers, 1632 mothers had EPDS scores <7, 177 had minimal English skills, and 150 mothers were not included as they were identified by nurses as having high levels of distress and, consistent with clinical practice guidelines, were referred for additional support [24]. There were a further 41 mothers who were identified by nurses as not being eligible for the study, but as a result of an administrative error, the reason for noneligibility was not provided to the research team. This left 213 mothers who scored >7 on the EPDS *and* reported at least 1 problem on the 4-item parenting problem questionnaire. A member of the research team telephoned these mothers to obtain verbal consent for a home visit by a study field worker. Among these mothers, 34 could not subsequently be contacted by telephone, 43 declined to participate in the study when contacted by telephone, and 3 were found to be ineligible. This left 133 mothers who completed the baseline assessment and were randomly assigned to the intervention or standard care arms of the study (Figure 1). Written consent was obtained at the time of this preintervention home visit.

Figure 1. Flow chart of participants.



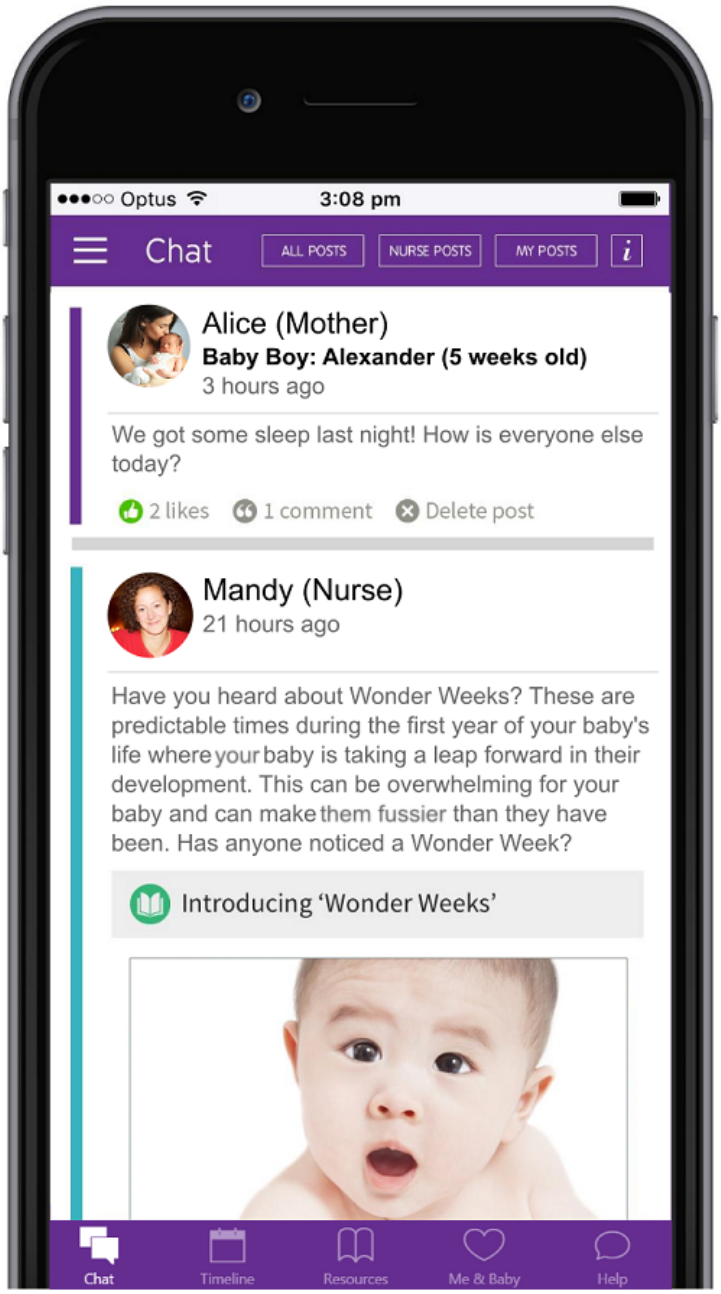
XSL•FO
RenderX

Full details of the research design are provided in the trial protocol [22]. In brief, the trial utilized a block randomized control design. Blocks of 20 consecutive mothers were randomized to either the intervention arm or the comparison (*standard care*) arm of the study using a randomization schedule that was generated by a statistician who was independent of the study team. The research team was blind to group allocation at the time of recruitment and assignment of mothers to the study groups. However, owing to the nature of the intervention, after the intervention commenced, it was not possible to keep the research staff or field workers blind to the groups to which mothers had been allocated. The exception to this was research staff coding the Nursing Child Assessment Satellite Training (NCAST) Parent-Child Interaction scale, who were blind to group allocation while completing coding. Further information regarding the randomization process is provided in the paper describing the trial protocol [22].

Intervention Versus Standard Care

Mothers randomized to the intervention arm of the study were assigned to a nurse-led, online group consisting of approximately 20 mothers of similarly aged infants. The 4-month intervention was delivered when infants were aged approximately 2 to 6 months and was accessed by mothers via a mobile phone app (see Figure 2). The app was free to download from the relevant mobile app store (ie, iTunes or Google Play), and mothers accessed the app on their own personal mobile devices (mobile phones or tablets). Nursing staff delivering the intervention were trained in the use and management of the app. They also received additional training in the mental health components of the intervention. [22]

Figure 2. The eMums Plus mobile phone app.



XSL•FO
RenderX

Full details of the intervention are provided in the manuscript describing the trial protocol [22]. New mothers and CaFHS nurses were involved in the development of the intervention. In addition, new mothers participating in CaFHS face-to-face mothers' groups were utilized to test the usability of the intervention app, with iterative improvements made before its use in the study, based on their advice.

Mothers in the comparison arm received *standard care*. In South Australia, standard care consists of a single home visit by a CaFHS nurse, usually within 4 weeks of the infant's birth. The nurse checks the health of the mother and baby, provides general advice about infant care, and provides information about community services available to help mothers and infants. Mothers in the intervention arm could also access community services if they wished to do this. Mothers in the intervention arm were not paid for their participation; however, a "thank you" gift was posted to the child of the participating mother upon completion of each questionnaire (3 in total).

Measures

Trained field workers administered self-complete questionnaires to mothers in their homes. Field workers also video-recorded mothers completing the NCAST teaching interaction during home visits. All measures, including the NCAST, were completed when infants were aged 1 to 2 months (preintervention), 8 months, and 12 months.

Primary Outcomes

Maternal Depressive Symptoms

The level of depressive symptoms experienced by mothers was assessed using the EPDS. The EPDS is a 10-item self-report questionnaire that assesses symptomatology during the previous 7 days. Scores range from 0 to 30 with higher scores indicating higher levels of depressive symptoms. Cut-off points can also be used to identify mothers who may benefit from additional support [23].

Maternal Self-Competence

1. Mothers' perceptions of their parental efficacy and their satisfaction with the parenting role were assessed using the 16-item Parenting Sense of Competence Scale (PSCS). Scores range from 16 to 96 with higher scores indicating higher levels of parenting self-competence and satisfaction with their parenting role [25,26].
2. The Parenting Stress Index (PSI) Competence subscale (excluding 2 items assessing parental education) was used to assess mothers' perception of their competence in caring for their infant [27]. The scale consisted of 11 items with scores ranging from 11 to 55. Higher scores indicated lower self-competence.

Quality of Mother-Infant Relationship

1. The NCAST Scale was used to assess the quality of mother-child interactions. The NCAST utilizes a 3 to 5-min video-recording of mothers teaching their child a skill appropriate to the age of their child. For the purpose of this study, we used the Teaching Scale as it is suitable for use with infants aged 0 to 36 months. Mother-child interactions are assessed in 6 areas: sensitivity to cues, response to distress, fostering social-emotional functioning, fostering cognitive growth, clarity of cues, and responsiveness to the caregiver [28]. Trained research assistants coded the video-recordings to generate a total score for mothers (including the subscales' sensitivity to cues, response to distress, fostering social-emotional functioning, and fostering cognitive growth) and a total score for infants (including the subscales' clarity of cues and responsiveness to the caregiver) [28]. Higher scores indicate higher levels of positive mother-child interaction quality. Further details about the NCAST scores are included in the Statistical Analyses section below.
2. The PSI Attachment subscale was used to assess mothers' perceptions of the quality of their relationship with their infant [27]. The scale consisted of 7 items with scores ranging from 7 to 35. Higher scores indicated the lower quality of mother-infant bonding.

The approach of using more than one measure to rate primary outcomes was employed to determine the consistency of the results, regardless of the particular measure used to assess the outcome. The analysis of results was undertaken independently for each measure.

Secondary Outcomes

1. Service utilization: Maternal self-completed questionnaires were used to identify other online or face-to-face health services used by mothers and infants.
2. Intervention quality: At the 8-month assessment, mothers completed a 40-item questionnaire designed by the researchers to assess mothers' perceptions of the quality of the intervention. Items in the questionnaire asked about the helpfulness of the intervention and the usability of the mobile phone app.
3. App usage: Data about app usage were automatically collected and included the number of log-ins, comments, and replies that mothers posted and the number of times mothers accessed the different elements of the intervention available on the app.

Ethics Approval

Ethics approval was received from the Women's and Children's Health Network Human Research Ethics Committee (approval numbers SSA/16/WCHN/016, HREC/16/WCHN/014).

Missing Data

In total, 22 participants did not complete the 8-month assessment (N=2), the 12-month assessment (N=7), or both assessments (N=13). We compared demographic characteristics and levels of functioning at baseline of these 22 participants and those with complete data at all assessments (N=111). Compared with mothers who completed all assessments, those who missed an assessment were (1) younger (missed assessment mean 28.4 years vs completed assessment mean 31.7 years; *95 % CI* −5.4 to −1.2); (2) had delivered their first child (missed assessment percentage=90.9% (20/22) vs completed assessment percentage=61.3% (68/111) ;P=.01); (3) less likely to have university qualifications (missed assessment percentage=40.9% (9/22) vs completed assessment percentage=64.9% (72/111) ;P=.02); and (4) had lower EPDS scores at their baseline

XSL•FO
RenderX

assessment (missed assessment mean 7.2 vs completed assessment mean 9.1; *95 % CI* −3.7 to −0.1).

Multiple imputations could not be performed owing to low cell sizes in the main predictors of missingness. We also attempted to calculate inverse probability response weights to account for the missing data, but weights could not be calculated owing to low cell sizes. As a result, we have presented complete case analyses in the main paper and analyses for all mothers with at least 1 outcome score at 8 or 12 months in the supplementary tables (see Multimedia Appendices 1 and 2) [29].

A small number of items were not completed on some measures. Where this occurred, missing item scores were estimated on the basis of mean item scores for the measure, as recommended by instructions for the measure.

Statistical Analyses

All analyses were intention-to-treat. Initially, unadjusted results were examined. Subsequently, adjusted mean scores for intervention and standard care groups at each time point were compared using linear generalized estimating equations) using exchangeable within-group correlation structures [30]. In each model, predictor variables were group (intervention and standard care), time (baseline, 8 months, and 12 months), and a group-by-time interaction. Analyses were adjusted for residual differences in demographic characteristics after randomization including the number of children, maternal education, housing situation, maternal age (years) at baseline, and the outcome score at baseline. All analyses were conducted using STATA 15.1 (StataCorp) [31].

There were 4 online mothers' groups in the study. However, intraclass correlations were very small at both 8- and 12-month assessments, ranging from 0.00 to 0.01 for the study outcomes. This indicated that there was little clustering of outcomes within groups. As a result, we did not adjust for clustering in the analysis of results.

A number of problems were identified with the NCAST videos recorded during the baseline and 8-month assessments. These included, for example, a failure to record the start of the teaching interactions; interruptions by field workers during teaching interactions (such as providing additional instructions or speaking to the child); and a failure to record the mother's or baby's face during some of the interaction tasks. These problems have the potential to invalidate NCAST scoring. After the problems were identified, further training was provided for field workers. As a result, the problems were largely overcome by the time of the 12-month assessment. However, given the problems with earlier assessments, we have only included results from the 12-month NCAST assessment in this study.

At the 12-month assessment, a small number of items were coded as *missing* for a large percentage of mothers because it was difficult to record the required mother-child behavior. As a result, items that could not be coded for more than 12% (13/113) of participants were not included in the NCAST score. In total, scores from 1 item were not included in the mother total score ("the caregiver provides non-verbal feedback to the child after the child performs better than the last attempt") and 4 items were not included in the child total score ("the child smiles/laughs during the episode", "the child grimaces/frowns during the episode", "the child smiles at caregiver within five seconds after caregiver's verbalization", and "the child smiles at caregiver within five seconds after caregiver's gesture, touch or facial expression"). As such, the mother total score consisted of 49 items (rather than 50) with possible scores ranging between 0 and 49. The child total score consisted of 19 items (rather than 23) with possible scores ranging between 0 and 19.

A linear regression analysis was used to compare the intervention versus standard care group NCAST scores for mothers and infants at the 12-month assessment. The scores were adjusted for the number of children, maternal education, housing situation, and maternal age (years) at baseline.

The frequency with which mothers logged on to the intervention, their frequency of message posting, and the extent to which they utilized different elements of the intervention are described using means and median scores. The number of additional services (both online and face-to-face services) that mothers used to support them in caring for themselves or their infants was assessed using percentage scores.

Sample Size

The sample size target for this study was 160 (80 in each trial arm). This sample size would provide .80 power to detect an effect size of Cohen *d*=0.4 at alpha=.05.

Results

The demographic characteristics and baseline scores for mothers in the complete case samples are shown in Table 1. The demographic characteristics and baseline scores for mothers in the response samples are shown in Multimedia Appendix 2.

Compared with mothers in the standard care group, mothers in the intervention group more frequently had more than 1 child, were younger, lived in rental accommodation, and had less frequently completed tertiary education. They also reported lower scores on the PSI competence subscale and higher scores on the PSCS and lower scores on the PSI attachment subscale (indicating better functioning on all 3 measures). A larger number of mothers in the intervention group than the standard care group could not be contacted at the time of follow-up assessments (Figure 1). It is possible that this was due to mothers in the intervention group being a more mobile population as reflected in their higher frequency of living in rental accommodation. All these postrandomization differences between the groups were adjusted for in the analysis of results.

The unadjusted outcome scores are presented in Table 2. Mothers in the intervention group had lower PSI competence scores and higher PSCS scores (indicating better functioning on both measures). They also had lower PSI attachment scores (also indicating better functioning). There was little difference in the EPDS scores between the intervention and standard care groups, except at the 12-month assessment where, on average, mothers in the standard care group had lower EPDS scores than mothers in the intervention group.

XSL•FO
RenderX

Table 1. Baseline demographic characteristics of children and mothers in the complete case samples in the intervention and standard care groups.

Characteristics	Complete case				*P* value[a]
	Intervention		Standard care		
	Values	n	Values	N	
First child, n (%)	29 (54)	54	39 (68)	57	.11
Male child, n (%)	29 (54)	54	27 (47)	57	.51
Child indigenous, n (%)	2 (4)	54	0 (0)	57	.14
Single parent household, n (%)	1 (2)	54	2 (4)	57	.59
Maternal age (years), mean (SD)	31.1 (5)	54	32.2 (4)	57	.16
Mother's education[b], n (%)					
University degree	28 (52)	54	44 (77)	57	.008
Trade or technical school	20 (37)	54	7 (12)	57	.008
Some or all years of high school	6 (11)	54	6 (11)	57	.008
Mother's employment, n (%)					
Full-time paid employment	32 (59)	54	36 (63)	57	.51
Part-time paid employment	16 (30)	54	18 (32)	57	.51
Other (self-employed or casual)	3 (6)	54	0 (0)	57	.51
Unemployed	3 (6)	54	3 (5)	57	.51
Housing, n (%)					
Rental or other	24 (44)	54	15 (26)	57	.046
Own home	30 (56)	54	42 (74)	57	.046
Currently breastfeeding, n (%)					
Yes	46 (85)	54	51 (90)	57	.5
No	8 (15)	54	6 (11)	57	.5
Partner's education[b,c], n (%)					
University degree	21 (40)	53	26 (47)	55	.63
Trade or technical school	20 (38)	53	20 (36)	55	.63
Some or all years of high school	12 (23)	53	9 (16)	55	.63
Partner's employment[c], n (%)					
Full-time paid employment	46 (87)	53	46 (84)	55	.49
Part-time paid employment	2 (4)	53	5 (9)	55	.49
Other (self-employed, contract, or casual)	3 (6)	53	4 (7)	55	.49
Unemployed	2 (4)	53	0 (0)	55	.49
Parenting stress index[d], mean (SD)					
Competence	26.5 (5.2)	54	28.8 (5.8)	57	.03
Attachment	10.9 (2.9)	54	13.6 (4.8)	57	.001
Maternal caregiving, mean (SD)					
Parenting Sense of Competence Scale	67.4 (10.3)	54	61.6 (9.7)	57	.003
Edinburgh Postnatal Depression Scale, mean (SD)	8.8 (3.0)	54	9.5 (4.9)	57	.38

[a]For the comparison between the intervention and standard care groups for the complete case sample.

[b]Highest level of completed education.

[c]Note that the single parents in the intervention and standard care conditions did not report their partner's education or employment; as a result, the complete case sample has fewer n values in these cells, but these parents are not excluded.

[d]For the Parenting Stress Index, higher scores indicate a worse outcome.

XSL•FO
RenderX

Table 2. Unadjusted mean outcome scores and mean difference (95% CI) between scores in the intervention and standard care groups.

Outcome assessment	Intervention (N=54)	Standard care (N=57)	Mean difference (95% CI)
Maternal confidence			
Parenting Stress Index (PSI) competence[a]			
Baseline	26.5	28.8	2.3 (0.2 to 4.4)
8 months	24.4	25.8	1.5 (−0.7 to 3.7)
12 months	23.6	25.2	1.6 (−0.5 to 3.8)
Parent Sense of Competence Scale			
Baseline	67.6	61.6	−5.7 (−9.5 to −1.9)
8 months	69.2	64.9	−4.4 (−7.9 to −0.9)
12 months	69.5	67.3	−2.2 (−5.9 to 1.6)
Relationship quality			
PSI attachment[a]			
Baseline	10.9	13.6	2.7 (1.1 to 4.2)
8 months	10.4	12.5	2.2 (0.9 to 3.5)
12 months	10.9	12.1	1.2 (−0.1 to 2.6)
Maternal Depression			
Edinburgh Postnatal Depression Scale			
Baseline	8.8	9.5	0.7 (−0.9 to 2.2)
8 months	7.9	8.7	0.7 (−1.1 to 2.5)
12 months	8.6	7.0	−1.5 (−3.2 to 0.1)

[a]Higher scores indicate more problems.

The adjusted outcome scores are shown in Table 3. There was little difference in the adjusted mean PSI competence scores between the groups at any of the assessments. However, at the baseline assessment, the adjusted mean Parent Sense of Competence score was higher for mothers in the intervention group (indicating better functioning) than for mothers in the standard care group. In contrast, there was little difference in the adjusted mean Parent Sense of Competence scores across the groups at the time of the 8-month and 12-month assessments.

The adjusted mean PSI attachment score was lower for mothers in the intervention group (indicating better quality relationships) than mothers in the standard care group at both the baseline and 8-month assessments but differed by a little across the groups at the 12-month assessment. There were only small differences in the adjusted mean EPDS scores between the groups at each assessment.

In the standard care group but not the intervention group, Parenting Sense of Competence scores were higher at the 12-month assessment than at baseline and EPDS scores were lower (indicating better functioning on each questionnaire).

A possible explanation for the improved functioning in the standard care group was that mothers in the group accessed more additional health services than mothers in the intervention group. We investigated this possibility by reviewing responses to questions asking about the frequency with which mothers in each group accessed face-to-face or online health services. There were few significant differences in service use between the groups but mothers in the intervention group had more frequently visited their family doctor and also a hospital emergency department *2 or more times* during the previous 6 months at the 12-month assessment (family doctor: intervention group=78% (42/54), standard care group=63% (36/57), *P*=.09; hospital emergency department: intervention group=15% (8/54), standard care group=4% (2/54), *P*=.04). They had also more frequently accessed online resources such as the *Baby Centre* website at the 8-month assessment (weekly/daily—intervention group=41% (22/54), standard care group=21% (12/56), *P*=.03) and the *Red nose* website or app at the 12-month assessment (monthly/weekly/daily—intervention group=67% (36/54), standard care group=46% (26/57), *P*=.03).

There was no difference in either the mother or infant adjusted mean NCAST total scores between the intervention and standard care groups (Table 4).

XSL•FO
RenderX

Table 3. Adjusted mean (95% CI) outcome scores in the intervention (N=54) and standard care (N=57) groups. All scores are adjusted for the number of children, maternal education, housing situation, and maternal age (years) at baseline.

Outcome assessment	Intervention, mean (95% CI)	Standard care, mean (95% CI)	Group × time, *P* value[a]
Maternal confidence			
Parenting Stress Index (PSI) competence[b]			
Baseline	26.5 (25.2 to 27.9)	28.8 (27.2 to 30.4)	.69
8 months	24.4 (22.7 to 26.1)	25.8 (24.4 to 27.2)	.69
12 months	23.6 (22.2 to 24.9)	25.2 (23.6 to 26.8)	.69
Parent Sense of Competence Scale			
Baseline	67.2 (64.5 to 69.9)	61.8 (59.1 to 64.4)	.11
8 months	69.1 (66.5 to 71.7)	64.9 (62.8 to 67.2)	.11
12 months	69.4 (66.8 to 72.0)	67.4 (64.7 to 70.1)	.11
Relationship quality			
PSI attachment[b]			
Baseline	10.9 (10.1 to 11.8)	13.6 (12.4 to 14.9)	.04
8 months	10.3 (9.4 to 11.2)	12.6 (11.6 to 13.5)	.04
12 months	10.9 (9.9 to 11.7)	12.1 (11.1 to 13.1)	.04
Maternal Depression			
Edinburgh Postnatal Depression Scale			
Baseline	8.6 (7.7 to 9.5)	9.6 (8.3 to 10.9)	.001
8 months	7.8 (6.6 to 9.0)	8.8 (7.5 to 10.1)	.001
12 months	8.4 (7.2 to 9.6)	7.2 (5.9 to 8.3)	.001

[a]Statistical significance of the group × time interaction.

[b]Higher scores indicate more problems.

Table 4. Adjusted Nursing Child Assessment Satellite Training mean (95% CI) scores for intervention and standard care groups in the complete case sample. All scores are adjusted for the number of children, maternal education, housing situation, and maternal age (years) at baseline.

NCAST[a] assessment at 12 months	Intervention		Standard care		Statistics, B[b] (95% CI)
	Mean (95% CI)	N	Mean (95% CI)	N	
Mother total score					
Unadjusted	36.5 (35.6 to 37.3)	51	36.2 (35.3 to 37.2)	54	—[c]
Adjusted	36.6 (35.7 to 37.5)	51	36.1 (35.2 to 37.0)	54	0.47 (−0.85 to 1.80)
Child total score					
Unadjusted	15.5 (15.1 to 15.9)	51	15.5 (14.9 to 15.9)	52	—
Adjusted	15.5 (15.0 to 15.9)	51	15.4 (14.9 to 15.9)	52	0.06 (−0.66 to 0.78)

[a]NCAST: Nursing Child Assessment Satellite Training; note that owing to administration errors, the n value varies slightly from the full complete case sample.

[b]Unstandardized regression coefficient.

[c]Not applicable.

Percentage of Mothers Logging Into the Intervention, Posting Messages, and Using Each Element of the Intervention

In the first 11 weeks of the intervention, more than 60% (43/72) of participants logged into the intervention at least once each week (the first table in Multimedia Appendix 3). Furthermore, more than 50% (38/72) of mothers logged into the intervention at least once each week until the 14th week of the 16-week intervention. This suggests that most mothers were regularly engaging with the intervention for the vast majority of the 16 weeks in which it was delivered. A somewhat lower percentage of mothers were regularly posting messages on the Chat page (the first table in Multimedia Appendix 3). However, almost

XSL•FO
RenderX

50% (36/72) of mothers posted at least 1 message each week until the 6th week of the 16-week intervention. Subsequent to this, the percentage of mothers posting at least 1 message each week varied, but even in the 10th week, 50% (36/72) of mothers posted at least 1 message.

The intervention app had a number of components that provided mothers with different kinds of information and support (see Multimedia Appendix 4). Full details about these elements are available in the paper describing the study protocol [22]. The second table in Multimedia Appendix 3 shows the percentage of mothers who accessed each component at least once per week. The Chat page where mothers could communicate with each other and where nurses posted comments and responded to queries was the component most frequently accessed (the second table in Multimedia Appendix 3). In contrast, after the initial weeks of the intervention, mothers less frequently accessed the other app components, including the resources component that contained key intervention content designed to support parenting and maternal emotional health (the second table in Multimedia Appendix 3).

Qualitative Assessment of the Intervention by Mothers

At the 8-month assessment (after they had completed the intervention), mothers were asked if they had received sufficient information to effectively use the app features. In total, 80% (47/59) of mothers *strongly agreed* or *agreed* that they had received sufficient information to use the Chat and Resources components of the app whereas 68% (40/59) to 76% (45/59) *strongly agreed* or *agreed* that they had received sufficient information to use each of the other app components.

At the 8-month assessment, mothers were also asked to rate the helpfulness and user-friendliness of each component of the intervention. In all areas, the majority of mothers reported that intervention components were *very helpful* or *helpful* (the third table in Multimedia Appendix 3). With the exception of mood graphing and video components of the app, the majority of mothers who used each component also reported that it was *very easy* or *easy* to use (the fourth table in Multimedia Appendix 3). However, a large percentage (30% (18/60) to 47% (28/60)) of mothers did not use some key components of the app such as the *mood-rater* designed to allow mothers to track their mood level over time and activities embedded in topic areas that were designed to help improve maternal emotional health (the fourth table in Multimedia Appendix 3). Finally, 90% (52/58) of mothers reported that the length of the information in each topic area was *about right*. In addition, 44% (25/57) of mothers reported the length of the intervention was *about right*, whereas 51% (29/57) reported that it was *too short*.

Discussion

Principal Findings

This study examined the effectiveness of a 4-month online group–based intervention led by community child health nurses and delivered via a mobile phone app. The intervention aimed to improve outcomes for mothers experiencing depressive symptoms and difficulty caring for their infants.

When infants were aged 8 and 12 months, maternal ratings of their level of maternal depressive symptoms and their parenting competence did not differ across the intervention and standard care groups. Standard care provided to mothers in each group consisted of a single home visit by a CaFHS nurse, within 4 weeks of the infant's birth, to check the health of the mother and baby and to provide information about relevant community services for mothers and infants. When infants were aged 12 months, there was also no difference between the observed quality of mother-infant interactions in the intervention versus standard care groups. The lack of differences between the groups occurred despite mothers in the intervention group engaging well with the intervention, logging into the app regularly, and posting comments more frequently than has been reported with previous online interventions [32,33]. In addition, a high percentage of mothers in the intervention group reported that the app was user-friendly, and its components were helpful for them.

There are 2 possible explanations for the findings from the study. First, although most mothers in the intervention group regularly logged into the app and communicated with each other through the Chat page, they made much less use of the app's text-based resources designed to provide support for depressive symptoms and guidance about how to solve common parenting problems. Furthermore, mothers also made little use of the activities in each topic area that were designed to reinforce skills being addressed by the topic. As such, it is possible that although most mothers reported that the intervention was helpful and the app easy to use, a failure to fully utilize key components in the resource meant that the intervention did not produce measurable improvements in maternal depressive symptoms and maternal parenting competence, beyond those achieved with standard care.

Second, despite mothers being randomly assigned to the intervention and standard care groups, the intervention group contained a higher proportion of socially disadvantaged mothers than the standard care group. For example, at the baseline assessment a higher percentage of mothers in the intervention group were living in rental accommodation and had lower levels of educational achievement than mothers in the standard care group. It is possible that the greater level of social disadvantage among mothers in the intervention group made them less likely to benefit from a largely self-directed online group–based intervention. It is also possible that these differences contributed to a higher level of attrition among mothers in the intervention group owing to the research team being unable to relocate a larger number of mothers in the intervention group than in the standard care group.

One of the key motivations for examining the effectiveness of nurse support and an intervention provided online through a mobile phone app is the limited availability of traditional face-to-face services to support the large number of mothers who experience difficulties during the postnatal period. The aim was to use innovative app-based technology to increase the number of mothers with depressive symptoms and parenting problems who can be provided with support by existing health services. Internet delivery has the potential to allow nurses to provide ongoing support services without the need to travel to

XSL•FO
RenderX

mothers' homes, reduce costs of *no-show* visits, and allow one nurse to have contact with many mothers during a single day. For mothers, internet delivery has the potential to enable access to reliable, evidence-based *just-in-time* information along with both professional and peer support without the need to attend fixed-time appointments in clinics that may be geographically distant from their homes.

The results of the study suggest that during the postnatal period, mothers will engage and utilize support provided through a mobile phone app. They also suggest that mothers find this method of receiving support to be helpful and user-friendly. As such, it appears that support for mothers during the postnatal period provided using mobile phone technology has the potential to be an important addition to existing services. However, based on results in this study, a continuing challenge is to find ways of getting mothers to engage with key elements of online interventions to ensure that they receive a sufficient *dosage* of the intervention components required to reduce depressive symptoms and improve parenting skills.

Strengths and Limitations

One of the main strengths of the study was its pragmatic nature. The RCT was conducted within a routine service setting using existing administrative infrastructure to allocate participants and existing clinical staff to deliver the intervention as part of normal service delivery. As a result, the intervention is able to be taken up as a part of routine delivery by the service and the results are those that we can expect if the service takes up this intervention. A potential limitation of this study is that residual differences in demographic characteristics between the intervention and standard care groups existed despite randomization [34]. The intervention group had characteristics associated with a higher level of disadvantage that may have led to experiences of greater adversity throughout the intervention that could have affected their levels of depression and mother-infant interaction quality independently of the intervention and in ways that differed to mothers in the standard care group.

Conclusions

Large numbers of women experience comorbid depressive symptoms and difficulties with caregiving during the postnatal period. However, there are limited health services available to support these women and their infants. In particular, there is a lack of services that provide combined support for both depression and caregiving difficulties during this period of time.

The eMums intervention did not reduce depressive symptoms nor improve maternal caregiving. Changes to outcome scores on the EPDS and PSI attachment scale suggested there was some improvement in these areas for mothers in the standard care group, largely after the intervention had ended. However, these changes were very small and unlikely to be of clinical significance.

Mothers in the intervention group reported that the intervention was helpful, and the app was described as easy to use. As such, this method of treatment delivery has the potential to support larger numbers of mothers than is possible using clinic-based face-to-face treatments. The major challenge will be to find ways to ensure that participating mothers fully engage with all the active intervention components required to reduce their depressive symptoms and improve their parenting skills.

Acknowledgments

This work was supported by a Channel 7 Children's Research Foundation Grant (161258) and a National Health and Medical Research Council (NHMRC) Centres of Research Excellence Grant (1099422). JWL and MS have funding for an NHMRC Centres of Research Excellence (1099422). ACPS is supported by funds from the Centres of Research Excellence. These researchers are independent of the funding bodies. The other authors have no financial relationships relevant to this study. The funders did not play a role in the design and conduct of the study; collection management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results for participants with outcome scores complete for a least one assessment postintervention.

[DOCX File, 16KB - jmir_v21i6e13689_app1.docx]

Multimedia Appendix 2

Baseline demographic characteristics of children and mothers for the response sample in the intervention and standard care groups.

[DOCX File, 18KB - jmir_v21i6e13689_app2.docx]

Multimedia Appendix 3

Use of the intervention and mothers' reports of the quality of different elements of the intervention.

XSL•FO
RenderX

[DOCX File, 23KB - jmir_v21i6e13689_app3.docx]

Multimedia Appendix 4

Key components of the eMums intervention app.

[DOCX File, 13KB - jmir_v21i6e13689_app4.docx]

Multimedia Appendix 5

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 3MB - jmir_v21i6e13689_app5.pdf]

References

1. Murray L. The impact of postnatal depression on infant development. J Child Psychol & Psychiat 1992 Mar;33(3):543-561. [doi: 10.1111/j.1469-7610.1992.tb00890.x]
2. Stein A, Gath DH, Bucher J, Bond A, Day A, Cooper PJ. The relationship between post-natal depression and mother-child interaction. Br J Psychiatry 1991;158(01):46-52. [doi: 10.1192/bjp.158.1.46]
3. Murray L, Arteche A, Fearon P, Halligan S, Goodyer I, Cooper P. Maternal postnatal depression and the development of depression in offspring up to 16 years of age. J Am Acad Child Adolesc Psychiatry 2011 May;50(5):460-470. [doi: 10.1016/j.jaac.2011.02.001]
4. Field T. Postpartum depression effects on early interactions, parenting, and safety practices: a review. Infant Behav Dev 2010 Feb;33(1):1-6 [FREE Full text] [doi: 10.1016/j.infbeh.2009.10.005] [Medline: 19962196]
5. Field T. Infants of depressed mothers. Infant Behav Dev 1995 Jan;18(1):1-13. [doi: 10.1016/0163-6383(95)90003-9]
6. Murray L, Cooper P, Fearon P. Parenting difficulties and postnatal depression: implications for primary healthcare assessment and intervention. Community Pract 2014 Nov;87(11):34-38. [Medline: 25612413]
7. National Institute for HealthCare Excellence. 2014. Antenatal and postnatal mental health: clinical management and service guidance URL: https://www.nice.org.uk/guidance/cg192 [accessed 2019-04-26] [WebCite Cache ID 77ulK9vqU]
8. National Institute for HealthCare Excellence. Postnatal care up to 8 weeks after birth URL: https://www.nice.org.uk/guidance/cg37 [accessed 2019-04-26] [WebCite Cache ID 77uleeAtn]
9. Buist A, Bilszta J, Milgrom J, Condon J, Speelman C, Hayes B. The National Postnatal Depression Program. 2001. The beyondblue National Postnatal Depression Program, Prevention and Early Intervention- Final Report URL: https://www.beyondblue.org.au/docs/default-source/8.-perinatal-documents/bw0075-report-beyondblue-national-research-program-vol2.pdf?sfvrsn=2 [accessed 2019-04-26] [WebCite Cache ID 77ulqNwGE]
10. Plantin L, Daneback K. Parenthood, information and support on the internet. A literature review of research on parents and professionals online. BMC Fam Pract 2009 May 18;10:34 [FREE Full text] [doi: 10.1186/1471-2296-10-34] [Medline: 19450251]
11. Doty JL, Dworkin J. Online social support for parents: a critical review. Marriage Fam Rev 2014 Mar 03;50(2):174-198. [doi: 10.1080/01494929.2013.834027]
12. Australian Bureau of Statistics. 2014. 8146.0 - Household Use of Information Technology, Australia, 2016-17 URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/8146.0 [accessed 2019-04-26] [WebCite Cache ID 77um5f8UX]
13. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [Medline: 12020305]
14. Buist AE, Austin MV, Hayes BA, Speelman C, Bilszta JL, Gemmill AW, et al. Postnatal mental health of women giving birth in Australia 2002-2004: findings from the beyondblue National Postnatal Depression Program. Aust N Z J Psychiatry 2008 Jan;42(1):66-73. [doi: 10.1080/00048670701732749] [Medline: 18058446]
15. Gavin NI, Gaynes BN, Lohr KN, Meltzer-Brody S, Gartlehner G, Swinson T. Perinatal depression: a systematic review of prevalence and incidence. Obstet Gynecol 2005 Nov;106(5 Pt 1):1071-1083. [doi: 10.1097/01.AOG.0000183597.31630.db] [Medline: 16260528]
16. Wisner KL, Sit DK, McShea MC, Rizzo DM, Zoretich RA, Hughes CL, et al. Onset timing, thoughts of self-harm, and diagnoses in postpartum women with screen-positive depression findings. JAMA Psychiatry 2013 May;70(5):490-498 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.87] [Medline: 23487258]
17. Henshaw C, Foreman D, Cox J. Postnatal blues: a risk factor for postnatal depression. J Psychosom Obstet Gynaecol 2004;25(3-4):267-272. [doi: 10.1080/01674820400024414]
18. Dennis CL, Coghlan M, Vigod S. Can we identify mothers at-risk for postpartum anxiety in the immediate postpartum period using the State-Trait Anxiety Inventory? J Affect Disord 2013 Sep 25;150(3):1217-1220. [doi: 10.1016/j.jad.2013.05.049] [Medline: 23764383]

XSL•FO
RenderX

19. Sawyer MG, Reece CE, Bowering K, Jeffs D, Sawyer AC, Mittinty M, et al. Nurse-moderated internet-based support for new mothers: non-inferiority, randomized controlled trial. J Med Internet Res 2017 Dec 24;19(7):e258 [FREE Full text] [doi: 10.2196/jmir.6839] [Medline: 28739559]
20. Sawyer MG, Reece CE, Bowering K, Jeffs D, Sawyer AC, Peters JD, et al. Usage, adherence and attrition: how new mothers engage with a nurse-moderated web-based intervention to support maternal and infant health. A 9-month observational study. BMJ Open 2016 Dec 05;6(8):e009967 [FREE Full text] [doi: 10.1136/bmjopen-2015-009967] [Medline: 27496227]
21. Sawyer AC, Lynch J, Bowering K, Jeffs D, Clark J, Mpundu-Kaambwa C, et al. An equivalence evaluation of a nurse-moderated group-based internet support program for new mothers versus standard care: a pragmatic preference randomised controlled trial. BMC Pediatr 2014 May 06;14:119 [FREE Full text] [doi: 10.1186/1471-2431-14-119] [Medline: 24886238]
22. Sawyer AC, Kaim AL, Reece CE, McDonald D, Le H, Clark J, et al. Evaluating the effectiveness of an app-based nurse-moderated program for new mothers with depression and parenting problems (eMums Plus): protocol for a pragmatic randomized controlled trial. JMIR Res Protoc 2019 Jan 16;8(1):e11549 [FREE Full text] [doi: 10.2196/11549] [Medline: 30664487]
23. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987 Jun;150:782-786. [Medline: 3651732]
24. beyondblue. Clinical practice guidelines for depression and related disorders - anxiety, bipolar disorder and puerperal psychosis - in the perinatal period. In: A guideline for primary care health professionals. Melbourne: beyondblue: the national depression initiative. Melbourne: beyondblue; 2011.
25. Johnston C, Mash EJ. A measure of parenting satisfaction and efficacy. J Clin Child Psychol 1989 Jun;18(2):167-175. [doi: 10.1207/s15374424jccp1802_8]
26. Rogers H, Matthews J. The parenting sense of competence scale: investigation of the factor structure, reliability, and validity for an Australian sample. Aust Psychol 2004;39(1):88-96. [doi: 10.1080/00050060410001660380]
27. Abidin R. Parenting Stress Index (PSI-4). 4th edition. Florida: Odessa: PAR; 2012.
28. Oxford M, Findlay D. NCAST Caregiver/Parent-Child Interaction Teaching Manual. 2nd edition. Seattle: NCAST Programs, University of Washington, School of Nursing; 2013.
29. Allison P. Missing Data. California: Sage Publications; 2001:978.
30. Dupont W. Statistical Modeling for Biomedical Researchers: A Simple Introduction to the Analysis of Complex Data. Cambridge, England: Cambridge University Press; 2009.
31. Stata Statistical Software: Release 15 program. : College Station, TX: StataCorp LP URL: https://www.stata.com/ [accessed 2019-04-26] [WebCite Cache ID 77unoq5AK]
32. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014 Feb 04;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966] [Medline: 24496109]
33. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
34. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. Soc Sci Med 2018 Dec;210:2-21 [FREE Full text] [doi: 10.1016/j.socscimed.2017.12.005] [Medline: 29331519]
35. Portal Australia. 2016. URL: http://www.portal.com.au/ [accessed 2019-04-26] [WebCite Cache ID 77uo8q6HG]

Abbreviations

CaFHS: Child and Family Health Service
EPDS: Edinburgh Postnatal Depression Scale
NCAST: Nursing Child Assessment Satellite Training
NHMRC: National Health and Medical Research Council
PSCS: Parenting Sense of Competence Scale
PSI: Parenting Stress Index
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by G Eysenbach; submitted 13.02.19; peer-reviewed by G Strudwick, T Stacey; comments to author 20.03.19; revised version received 27.03.19; accepted 30.03.19; published 04.06.19.

Please cite as:
Sawyer A, Kaim A, Le HN, McDonald D, Mittinty M, Lynch J, Sawyer M
The Effectiveness of an App-Based Nurse-Moderated Program for New Mothers With Depression and Parenting Problems (eMums Plus): Pragmatic Randomized Controlled Trial
J Med Internet Res 2019;21(6):e13689
URL: https://www.jmir.org/2019/6/e13689/
doi:10.2196/13689
PMID:31165715

©Alyssa Sawyer, Amy Kaim, Huynh-Nhu Le, Denise McDonald, Murthy Mittinty, John Lynch, Michael Sawyer. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Outcomes and Device Usage for Fully Automated Internet Interventions Designed for a Smartphone or Personal Computer: The MobileQuit Smoking Cessation Randomized Controlled Trial

Brian G Danaher[1,2], PhD; Milagra S Tyler[1,2], MS; Ryann C Crowley[2,3], MS; Håvar Brendryen[4], PhD; John R Seeley[1,2], PhD

[1]Prevention Science Institute, University of Oregon, Eugene, OR, United States

[2]Oregon Research Institute, Eugene, OR, United States

[3]Center for Digital Mental Health, University of Oregon, Eugene, OR, United States

[4]Norwegian Centre for Addiction Research, University of Oslo, Oslo, Norway

Corresponding Author:
Brian G Danaher, PhD
Prevention Science Institute
University of Oregon
1600 Millrace Dr
Eugene, OR, 97403
United States
Phone: 1 5413464490
Email: bdanaher@uoregon.edu

Abstract

Background: Many best practice smoking cessation programs use fully automated internet interventions designed for nonmobile personal computers (desktop computers, laptops, and tablets). A relatively small number of smoking cessation interventions have been designed specifically for mobile devices such as smartphones.

Objective: This study examined the efficacy and usage patterns of two internet-based best practices smoking cessation interventions.

Methods: Overall, 1271 smokers who wanted to quit were randomly assigned to (1) MobileQuit (designed for—and constrained its use to—mobile devices, included text messaging, and embodied tunnel information architecture) or (2) QuitOnline (designed for nonmobile desktop or tablet computers, did not include text messages, and used a flexible hybrid matrix-hierarchical information architecture). Primary outcomes included self-reported 7-day point-prevalence smoking abstinence at 3- and 6-month follow-up assessments. Program visits were unobtrusively assessed (frequency, duration, and device used for access).

Results: Significantly more MobileQuit participants than QuitOnline participants reported quitting smoking. Abstinence rates using intention-to-treat analysis were 20.7% (131/633) vs 11.4% (73/638) at 3 months, 24.6% (156/633) vs 19.3% (123/638) at 6 months, and 15.8% (100/633) vs 8.8% (56/638) for both 3 and 6 months. Using Complete Cases, MobileQuit's advantage was significant at 3 months (45.6% [131/287] vs 28.4% [73/257]) and the combined 3 and 6 months (40.5% [100/247] vs 25.9% [56/216]) but not at 6 months (43.5% [156/359] vs 34.4% [123/329]). Participants in both conditions reported their program was usable and helpful. MobileQuit participants visited their program 5 times more frequently than did QuitOnline participants. Consistent with the MobileQuit's built-in constraint, 89.46% (8820/9859) of its visits were made on an intended mobile device, whereas 47.72% (691/1448) of visits to QuitOnline used an intended nonmobile device. Among MobileQuit participants, 76.0% (459/604) used only an intended mobile device, 23.0% (139/604) used both mobile and nonmobile devices, and 0.1% (6/604) used only a nonmobile device. Among QuitOnline participants, 31.3% (137/438) used only the intended nonmobile devices, 16.7% (73/438) used both mobile and nonmobile devices, and 52.1% (228/438) used only mobile devices (primarily smartphones).

Conclusions: This study provides evidence for optimizing intervention design for smartphones over a usual care internet approach in which interventions are designed primarily for use on nonmobile devices such as desktop computers, laptops. or tablets. We propose that future internet interventions should be designed for use on all of the devices (multiple screens) that users prefer. We forecast that the approach of designing internet interventions for mobile vs nonmobile devices will be replaced by internet interventions that use a single Web app designed to be responsive (adapt to different screen sizes and operating systems), share user data across devices, embody a pervasive information architecture, and complemented by text message notifications.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT01952236; https://clinicaltrials.gov/ct2/show/NCT01952236 (Archived by WebCite at http://www.webcitation.org/6zdSxqbf8)

(J Med Internet Res 2019;21(6):e13290) doi:10.2196/13290

KEYWORDS

tobacco; smoking; internet; eHealth; mHealth; smartphone; device

Introduction

Background

Many current best practice smoking cessation programs use fully automated internet interventions designed for personal computers (nonmobile devices such as desktop computers, laptops, and tablets) that provide media-rich, multifaceted content [1-7]. Owing to their substantial reach via the internet, these interventions offer the promise of helping large number of smokers who want to quit [8-11]. However, benefits derived from these internet interventions are probably reduced because they are delivered largely for personal computers that are not readily accessible during a user's everyday routine. Moreover, the interventions typically expect users to take the initiative to access the program. In contrast, just-in-time mobile internet interventions allow users to take the intervention with them during their everyday routines [12,13]. Mobile interventions take the initiative to proactively send or *push* content to users, including program reminders, strategy refreshers, and encouraging text messages [1,2,10,14-20]. Although mobile health interventions introduce new opportunities, they also come with some limitations. For example, the relatively smaller screens may require adaptations from traditional Web content in terms of shorter text and simpler graphics. A relatively small number of smoking cessation interventions reported in the research literature have been designed specifically for mobile devices [16], and to our knowledge, there are no direct comparisons of interventions designed for smartphones (mobile devices) vs interventions designed for desktop computers (nonmobile devices) previously reported. Finally, the context for this discussion is that most US adults own multiple *information devices*: (almost 77% use a smartphone, 75% use a desktop or laptop computer, and almost 50% use tablets [21]), and they use these multiple devices sequentially as well as at the same time [22].

Aims of This Research

This study examined the efficacy and usage patterns (including devices used to visit) of 2 internet interventions for smoking cessation both of which used best practice tobacco cessation content. The MobileQuit intervention was optimized for smartphones, whereas the QuitOnline intervention represented a usual care internet intervention in that it was designed primarily for use on nonmobile PCs (desktop, laptop, or tablet computers).

Methods

Participants Recruitment/Enrollment

A nationwide internet-based marketing campaign used Google AdWords, Reddit, Smokefree.gov, and ORI.org. Respondents completed an internet-based registration procedure (screening survey, steps validating a functional email account and a cellphone number, informed consent, contact information, and baseline assessment) before being assigned to condition via computer-generated randomization (not personal preference; see Multimedia Appendix 1). The study protocol was approved by the ORI Human Subjects Institutional Review Board (Assurance Identification #FWA00005934).

The eligibility criteria were as follows: (1) aged ≥18 years, (2) cigarettes were the primary tobacco product, (3) smoked ≥5 cigarettes/day for the previous 6 months, (4) smoked in the last 7 days, (5) wanted to quit smoking in next 14 days, (6) active users of a smartphone (iPhone or Android) and a personal computer or tablet, (7) willing to receive up to 150 text messages over 6 months of the program, (8) able to access the internet, (9) not have another household member participating in the research project, (10) have a valid personal email address, (11) have a valid mobile phone, and (12) US resident.

Tailored Welcome Messaging

Each participant received a welcome message announcing their treatment program assignment (Textbox 1). This message was tailored (with emphasis added) based on the treatment assignment and the type of device the participant used during screening.

XSL•FO
RenderX

Textbox 1. Tailored welcome messages.

If randomized to QuitOnline and the device being used at screening is a smartphone:

- *Congratulations you have been assigned to the stop smoking program designed especially for you to use on your desktop or tablet. We have sent you an email to confirm your participation and to help you get to your program from your desktop.*

If randomized to QuitOnline and the device at screening is not a smartphone:

- *Congratulations, you have been assigned to the QuitOnline program designed especially for you to use on your desktop or tablet. Please click on the Get Started button start using the program. We have also sent you an email to confirm your participation, and so you can get back to the program anytime you want.*

If randomized to MobileQuit and device at screening is a smartphone:

- *Congratulations, you have been assigned to the stop smoking program designed especially for you to use on your smartphone. Please click on the Get Started button to start using the program. We have also sent you an email to confirm your participation, and so you can get back to the program anytime you want.*

If randomized to MobileQuit and device at screening is not a smartphone:

- *Congratulations, you have been assigned to the stop smoking program designed especially for you to use on your smartphone. We have sent you an email to confirm your participation, and so you can get to your program from your smartphone.*

Intervention Conditions

The 2 internet interventions presented very similar best practice smoking cessation content based on cognitive behavior therapy features (see Table 1) including many of the same interactive and multimedia features (Table 2). Both emphasized the phases of quitting—Preparing to Quit, Quitting, Maintaining Abstinence, and Retooling if lapse.

Table 1. Cognitive behavior therapy ingredients in both internet interventions.

Cognitive behavior therapy ingredients	Features	Example
Explanation of the treatment model[a]	Display text and animation and frequently asked questions (FAQs).	Overview of preparing to quit, quitting, and maintaining nonsmoking.
Goal setting[a,b,c]	Display text, assign stars to list of choices to choose which strategies to use, and narrow choice via series of questions.	Set goals to quit smoking and maintain nonsmoking.
Tracking[b]	Periodic notification messages asking user to reply and view summary charts of key ratings.	Track smoking/nonsmoking status and track temptation (high smoking urge) situations.
Pleasant activities[a,b]	Display text, identify activities using a *list activity* that permits typing description of activity or choose from prepopulated list items.	Identify and plan for situations that trigger smoking urges.
Self-defeating thoughts[b]	Display text and FAQs, view animations showing procedures to identify and interrupt downward spirals, and videos of coping models.	Identify and interrupt downward spirals that lead to smoking.
Positive thoughts[b]	Display text and FAQs and videos of coping models.	Focus on being smokefree.
Stress management[b]	Display text and FAQs and videos of coping models.	Two brief relaxation strategies.
Maintenance plan[b]	Choose strategies to use and sign commitment contract.	Personal plan to maintain nonsmoking.
Relapse plan[a,b]	Review circumstances of lapse, list what to do differently, and sign commitment contract.	Plan for smoking slips.

[a]Increasing awareness (destigmatizing/normalizing).

[b]Providing opportunities for corrective experiences.

[c]Encouraging repeated practice.

XSL•FO
RenderX

Table 2. Participant engagement activities in both internet interventions.

Activities	Functions	Examples
List activities	Encourage creation of personal lists.	Lists of pleasant activities, list of supporters, reasons for wanting to feel better, contributing factors, high-tension situations, and warning signs.
Expand-collapse (accordion) activities	Enable exploration of additional detail on topics of interest.	Frequently asked questions, myths and facts.
Drag and drop activity	Provide interactive experience to test discrimination.	Differences between extreme thoughts and everyday concerns.
Goal-setting activity	Interactive steps for selecting goals.	Number of pleasant activities to accomplish each day and the strategies that worked.
Practice change activities	Doing homework tasks in normal routine.	Identify a downward spiral, practice relaxation, and anticipate and savor activities.
Behavior tracking	Chart data over time to identify patterns and show progress.	Daily tracking of smoking status plotted in a chart.
Host videos	Provide *human touch* and highlight topics in each session.	Host videos at the start of each session.
Testimonial videos	Coping models overcome challenges to quit smoking using strategies from program.	Other smokers' experiences, for example, doing more fun activities and managing mood patterns and stress.
Animated tutorials	Provide explanation for underlying models for change.	Show downward mood spiral and how it can be caught and managed at critical choice points.

MobileQuit Condition (Designed for Smartphone Delivery)

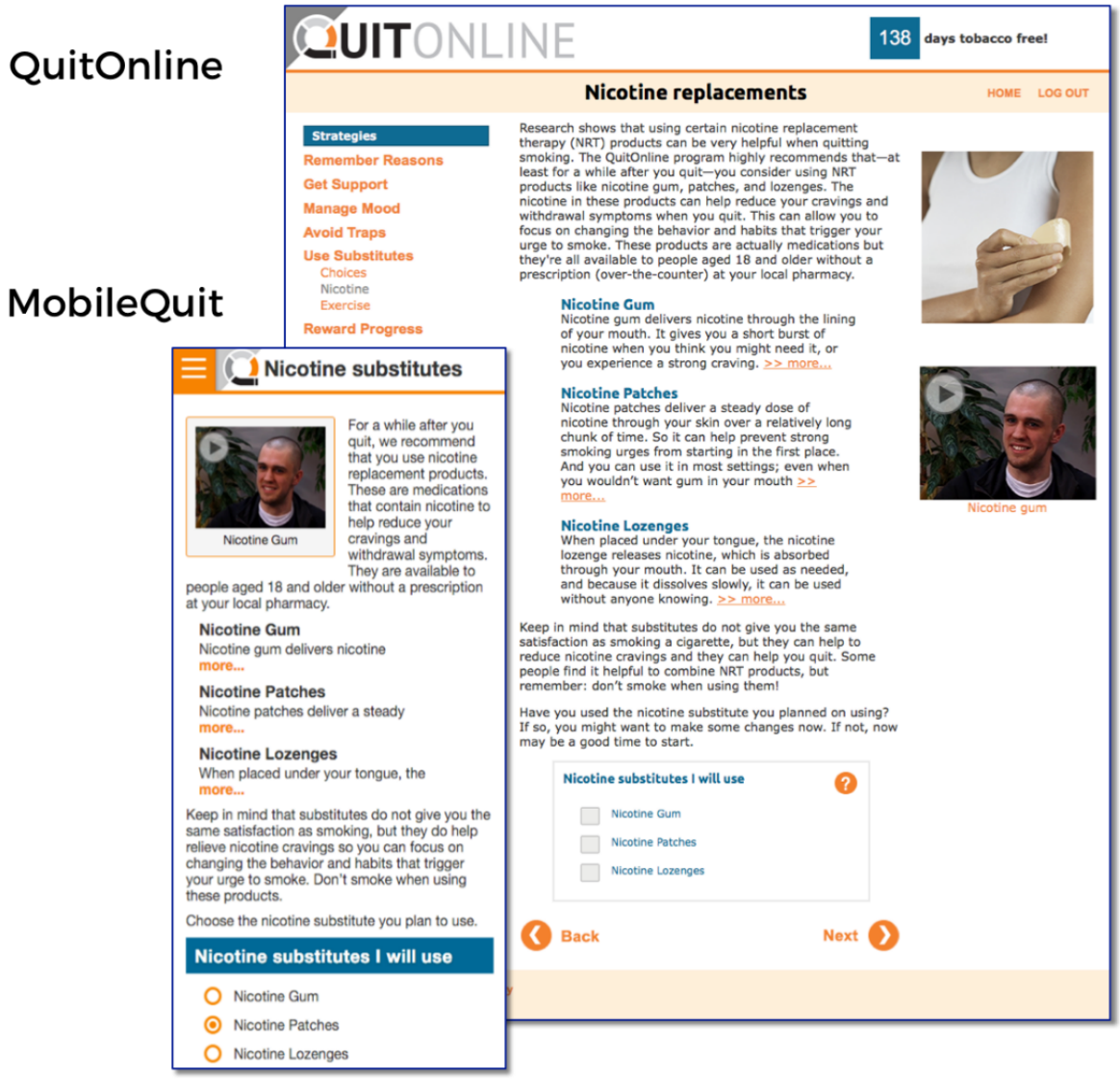
The MobileQuit condition used an integrated mobile Web app and text messaging intervention designed for a smartphone's Web browser and had an appearance and functionality similar to what would be found on a native app (eg, button on desktop for launch). Web apps are relatively uncomplicated to update, they use similar designs and programming across iOS and Android operating systems, and they permit unobtrusive monitoring of program usage [23]. As we wanted to examine differences by device type, our log-in system attempted to constrain access so that only smartphones could be used to access the MobileQuit program.

Information Architecture

MobileQuit used a tunnel information architecture [23-25] that defined the step-by-step order in which the program was delivered over time, similar to the one used by Brendryen et al [1,2] in their efficacious *Happy Ending* smoking cessation projects. Major *Topics of the Day* could be viewed for a single day, and then excerpts were available as an ongoing reference in the program's Library and Action Plan. Examples are displayed in Figures 1-4.

XSL•FO
RenderX

Figure 1. Screenshot 1 of MobileQuit and QuitOnline.

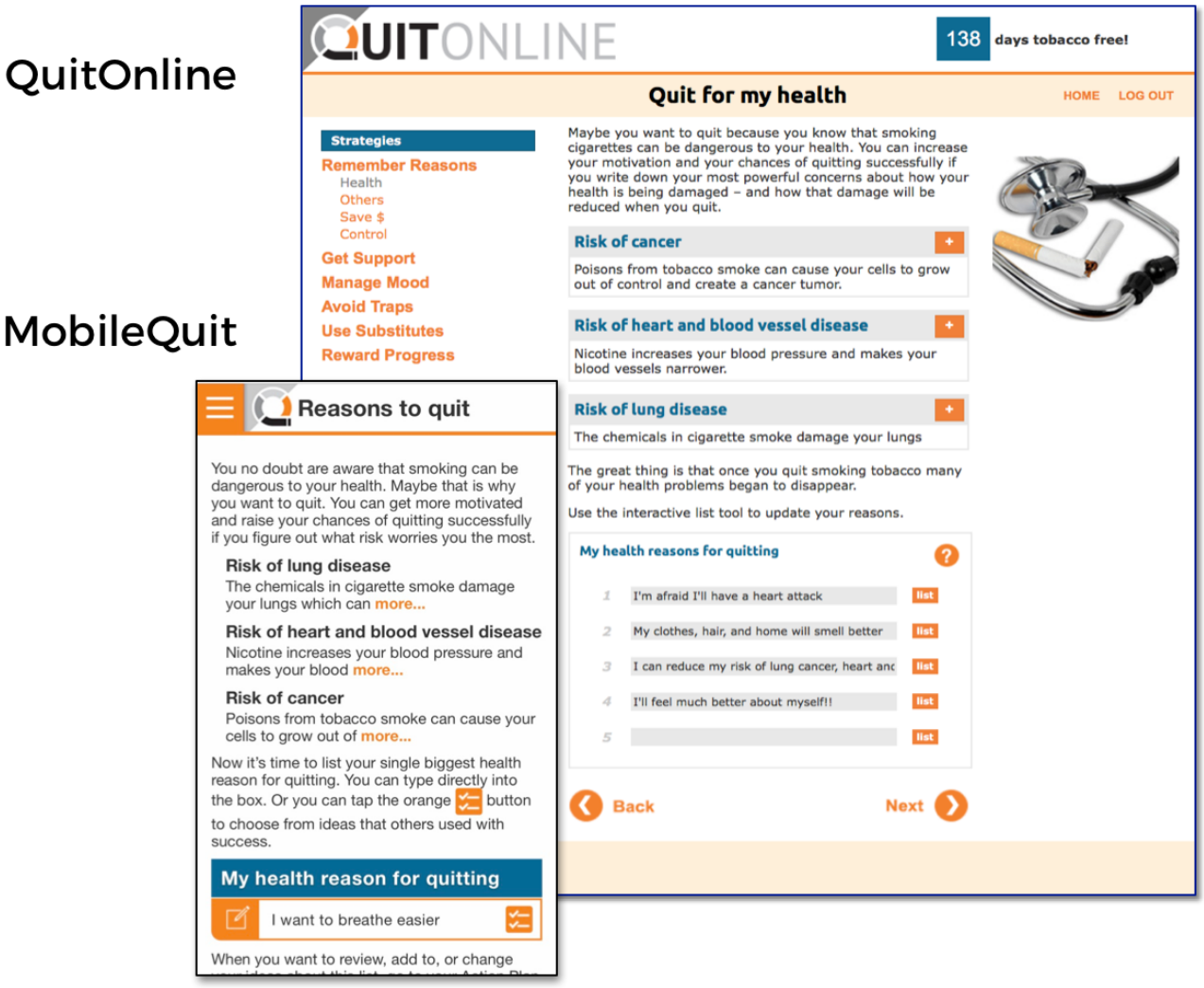


XSL•FO
RenderX

Figure 2. Screenshot 2 of MobileQuit and QuitOnline.

XSL•FO
RenderX

Figure 3. Screenshot 3 of MobileQuit and QuitOnline.



XSL•FO
RenderX

Figure 4. Screenshot 4 of MobileQuit and QuitOnline.

Lapse Management Using the Detective Activity

Participants who reported experiencing a lapse were encouraged to use the program's *Detective Activity* —an interactive wizard that asked a series of questions to help elicit the circumstances of the slip (Figure 5) to create a personal lapse-prevention plan. Participants could use the Detective Activity multiple times.

Text Message Content and Schedule

Participants were sent text messages that synchronized with their program's predefined tunnel schedule. As shown in Figure 6, a total of 290 text messages composed of 4 types of content were scheduled over the 6-month study period. Additional text messages were sent if the participant did not view certain program content, did not quit on the quit date, reported a lapse, reset the program's clock, replied to smoking status texts, or was scheduled for a follow-up assessment. Participants could opt out of receiving text messages at any time without dropping out of the study.

QuitOnline Condition (Designed for Desktop, Laptop, and Tablet)

The QuitOnline personal computer condition was an internet intervention that used interactive and multimedia components to deliver best practice smoking cessation content. Adapted from the efficacious MyLastDip smokeless tobacco cessation program [26,27], QuitOnline used a hybrid matrix-hierarchical information architecture [24] that enabled participants to freely examine available content. Participants were sent automated email reminders to visit their program following periods of inactivity or when they set a quit date.

Although intended for use on desktop computers, QuitOnline adjusted its functionality somewhat when used on a tablet to enable touch control, entering/editing text, and playing videos. It did not automatically adjust its appearance to fit the smaller screens of mobile devices.

Usability Testing

Both single-session and longitudinal usability testing methods were used. During single-sessions, usability testers (N=6; as recommended by Nielsen [28]) met in a research laboratory with a trained research staff member while interacting with the program and using the think-aloud procedure [29]. Consistent with use cases in usability testing [30,31] and experience sampling methods [32,33], testers followed the longitudinal usability approach that asked them to be engaged with the program during their normal routine over several weeks while keeping detailed notes. Example use cases for MobileQuit included not answering, quitting early, lapsing, and answering 2-way text messages. Testers also completed structured interviews at the end of the test period.

XSL•FO
RenderX

Figure 5. MobileQuit's detective activity.

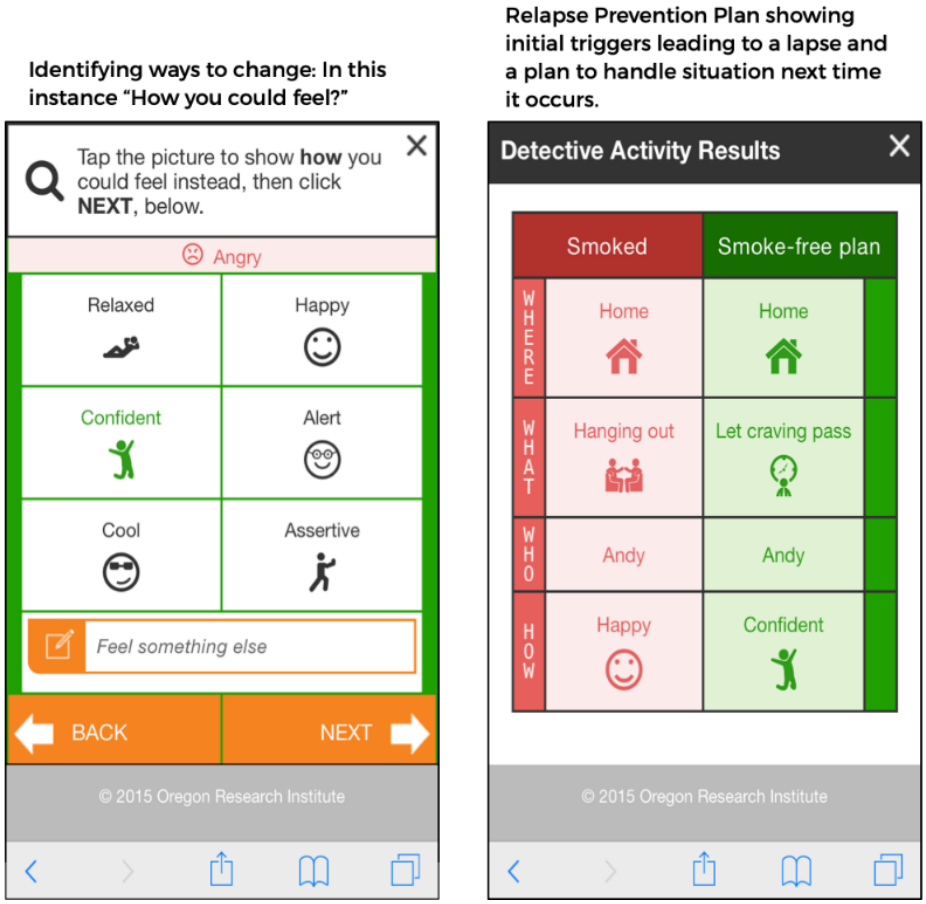
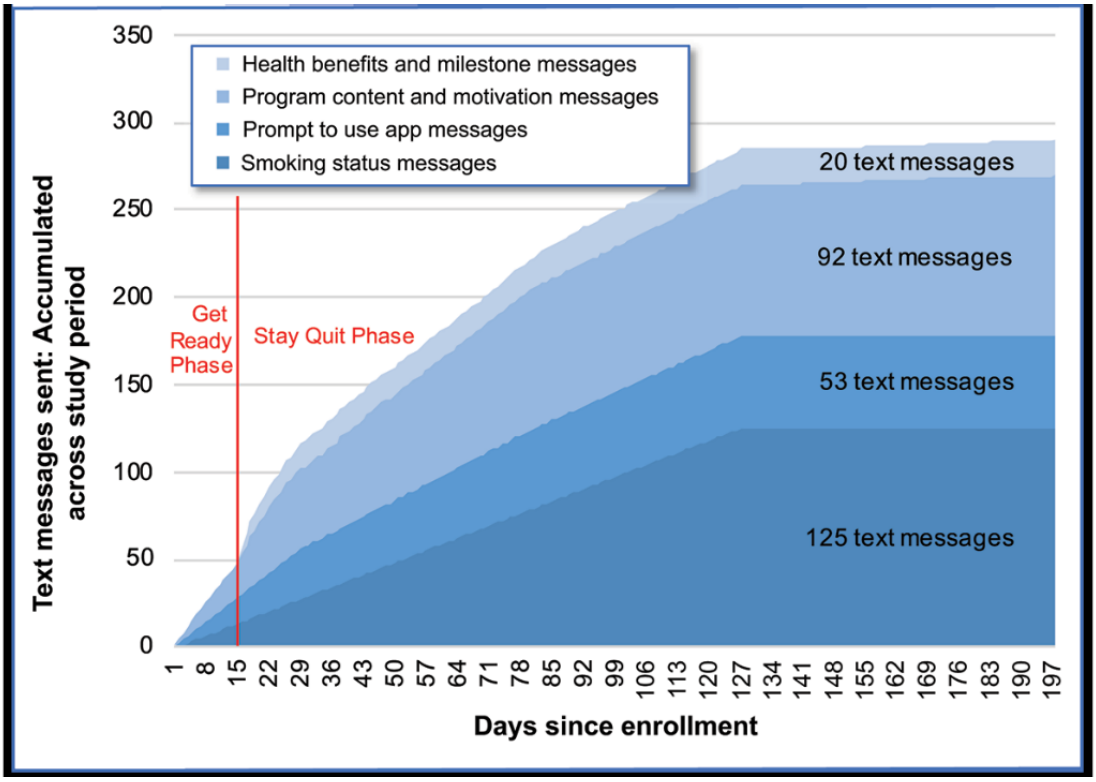


Figure 6. Standard regimen of 290 text messages planned to be sent to MobileQuit participants by message type.



XSL•FO
RenderX

Table 3. Schedule of assessments

Assessments	Screening and Baseline	3-month assessment	6-month assessment
Socio-demographics	X[a]	—[b]	—
Past tobacco use	X	—	—
Current tobacco use	X	X	X
Quit smoking status[c]	X	X	X
Nicotine dependence	X	—	—
Self-efficacy	X	X	X
Readiness to quit	X	—	—
Depression status	X	X	X
Alcohol use	X	—	—
Cannabis use	X	—	—
Helpfulness, usability and satisfaction	—	X	—
Use of other treatments	X	X	X
Device used to access program[d]	X	X	X
Use of program content[d]	X	X	X

[a]Indicates when the assessment occurred.

[b]Not applicable.

[c]Measured via texts and return user questions from enrollment through 6-month assessment.

[d]Measured continuously and unobtrusively from enrollment through 6-month assessment.

Assessment Plan and Measures

Baseline assessment was completed before randomization, and participants were sent an email reminder with the URL to encourage completion of follow-up assessments at 3 and 6 months (see Table 3). If a follow-up assessment was not completed after 2 weeks, research staff attempted to complete an assessment by phone. Any participant who did not complete a follow-up assessment within 45 days of its scheduled date was determined to have failed to complete that assessment. Participants received US $20 for each completed follow-up assessment and an additional US $20 if they completed both assessments. Remuneration was not tied to quitting smoking.

Sociodemographics

Data were collected on participant age, gender, race/ethnicity, marital status or long-term romantic relationship with a partner, and educational background.

Internet Usage

At both screening and baseline, the participants were asked about how they accessed the internet. For example, eligibility was determined in part by self-reported use of smartphones as well as other types of computers to access the internet. In addition, a baseline question was asked: “Overall, when you use the internet, do you do that mostly using your smartphone or mostly using some other device like a desktop, laptop or tablet computer?” Answer options included mostly on smartphone, mostly on something else, both equally, depends, or don’t know.

Current Tobacco Use

At screening, the respondents were asked about the number of cigarettes they smoked. Point prevalence self-reported smoking status was asked on all assessments: “In the past 7 days, have you smoked any cigarettes?” with answer options: no, not even a puff (scored 0) or yes (scored 1). If they reported that they had smoked, then they were asked “On average, how many cigarettes do you smoke in a day?” At follow-up, participants were asked “Since you enrolled in [assigned treatment program], when did you last smoke a cigarette?” with answer options less than 1 month ago, 1 month ago, 2 months ago, and 3 months ago.

Use of Other Tobacco Products, Quit Aids, and Nonassigned Treatments

The participants were asked about their use of any tobacco products other than cigarettes: “What type of tobacco products have you used in the past 7 days?” with answer options of E-cigarettes, cigars, pipe, chew/snuff, other [open ended text permitted], or I do not use any other tobacco products. At all 3 assessments participants were also asked: “Are you currently using any of the following to help you quit smoking?” Answer options (check all that apply) included: nicotine replacement therapy (NRT), patches, lozenges, or gum; prescription medication, such as bupropion (brand name Zyban, Wellbutrin) or varenicline (brand name Chantix); formal treatment (telephone quitlines, therapy including group and individual, hypnosis, acupuncture, etc); and none of the above.

XSL•FO
RenderX

Past Tobacco Use

At baseline, the participants were asked about their tobacco history (years of use, number of quit attempts, and amount of use) as well as smoking by spouse/partner, by household members, and among their 5 best friends. They were also asked: "How many times have you made a serious attempt to quit smoking cigarettes for more than 24 hours in the last 3 months?"

Slip Plans

Participants were asked the extent to which they endorsed a series of statements at baseline and the 3-month assessment: "I expect that I might slip and smoke a cigarette"; "Even if I slip, I still expect to quit smoking for good"; and "If I slip, I have a plan to get back on track to being smoke free." Each statement used the same endorsement options: strongly disagree (scored 0), disagree (scored 1), neither agree nor disagree (scored 2), agree (scored 3), and strongly agree (scored 4).

Nicotine Dependence

The 6-item Fagerstrom Test for Nicotine Dependence [34,35] was assessed at baseline. We separately examined the time of first smoke in the morning as this dependence item has been found to be highly predictive of subsequent abstinence [36].

Self-Efficacy

Participant self-confidence in quitting was assessed at each assessment by asking: "How confident are you that you will not be using tobacco a year from now" using a 5-point scale: not at all, a little, somewhat, moderately, and extremely. This item was used in our previous research [37] and was found to be a mediator of tobacco abstinence.

Readiness to Quit

The participants were asked at baseline to rate their confidence in quitting using the contemplation ladder [38] that has an 11-point scale with the following answer options: I have no thought about quitting smoking (scored 0), I think I need to consider quitting smoking someday (scored 2), should quit but not quite ready (scored 5), I am starting to think about how to reduce the number of cigarettes I smoke a day (scored 8), and I am taking action to quit smoking (scored 10).

Depressive Symptoms

Participant depressive symptomatology was assessed at baseline and at both follow-up assessments using the Patient Health Questionnaire-8 (PHQ-8) [39], which provides a validated measure of depression severity. The PHQ-8 asks: "Over the last 2 weeks, how often have you been bothered by any of the following problems?" with answer options of: not at all (scored 0), several days (scored 1), more than half the days (scored 2), and nearly every day (scored 3) [40,41]. A PHQ-8 score ≥10 has been found to have an 88% sensitivity and 88% specificity for major depression and typically represents clinically significant depression [42].

Alcohol Use

Alcohol use was assessed at baseline using a single item that asked, "On average during a typical week, how many drinks of alcohol do you have?" Heavy use was defined as ≥13 drinks/week for men and ≥7 drinks/week for women.

Cannabis Use

The 7-day point prevalence use of cannabis was assessed at baseline and the 3-month assessment using the question: "In the past 7 days, have you smoked cannabis (marijuana)?" which used dichotomous answer options of no, not even a puff (scored 0) or yes (scored 1).

Usability, Helpfulness, and Satisfaction

At the 3-month assessment, the participants were asked 2 questions about program usability and helpfulness:

"How easy was it for you to use the [MobileQuit; QuitOnline] program?" and "How helpful was the [MobileQuit; QuitOnline] program?" with answer options of not at all (scored 0), a little (scored 1), somewhat (scored 2), moderately (scored 3), and extremely (scored 4).

At the 3-month assessment, the participants in the MobileQuit condition were also asked questions about text messaging:

- "Did the MobileQuit text messages make it easier for you to quit?" with answer options of not at all (scored 0), a little (scored 1), somewhat (scored 2), moderately (scored 3), and extremely (scored 4);
- "How would you describe the number of text messages you received from MobileQuit?" using answer options of not enough, just the right number, too many, and no opinion.
- "Overall, what percentage of MobileQuit's text messages did you read?"

The participants were also asked at 3 months about their satisfaction with their assigned program: "Would you recommend the [MobileQuit; QuitOnline] program to friends or family members who are interested in quitting smoking?" with answer options of yes (scored 1), no (scored 0), and not sure (scored 2).

Participant Engagement (Use of Assigned Treatment)

Both interventions unobtrusively tracked the overall number and duration of website visits from enrollment to the end of the 6-month follow-up assessment [37,43]. Visits were required to last at least 1 second to be counted, and there could be multiple visits/day. The date/time of each text message was logged automatically by the program, although it was not technically possible to determine whether the participant viewed or read a text message or for how long.

Device Used to Access the Program

The device used by each participant to make each program visit was assessed unobtrusively using the ScientiaMobile Wireless Universal Resource FiLe (WURFL) tool that analyzed the user agent string sent by the browser [44,45]. Consistent with Google's method of categorizing mobile vs nonmobile devices [22,46], we considered mobile devices to include smartphones and feature phones whereas nonmobile devices included personal computers (desktop computers), laptops, and tablets.

Statistical Analyses

The results were analyzed separately for the 3- and 6-month follow-up assessments as well as using a repeated point prevalence measure that combined 3- and 6-month assessments

XSL•FO
RenderX

as a measure of more lasting abstinence. Logistic regression models were used to calculate the odds ratios for abstinence rate differences between intervention conditions, adjusting for significant baseline differences between conditions. Secondary analyses, assessing changes in cigarette usage (number of cigarettes per day) and quit attempts among participants who continued to use cigarettes, were analyzed using regression models with a covariate adjustment for baseline values.

The possible predictors of outcomes were assessed using a 2-step procedure. First, a univariate binary logistic regression was used to test the baseline participant characteristics as predictors of smoking abstinence using the repeated point abstinence at 3 and 6 months. Next, the significant predictors were tested using multivariate binary logistic regression with backward elimination. To identify any differential effects of intervention on outcomes, the multivariate test included treatment condition as well as the interaction of condition with sample characteristics.

IBM SPSS (version 24) was used for all statistical analyses, unless otherwise noted. Analyses used both intention-to-treat (ITT; in which participants who did not complete their assessments were considered to be using tobacco [47]) and Complete Cases (limited to participants who completed assessments). For the ITT analysis, there was sufficient power (.80) to detect a smoking abstinence rate difference of ≥7% between intervention conditions with alpha set to .017 (.05/3) to adjust for the 3 primary outcomes (3-month, 6-month, and repeated 3- and 6-month point prevalence rates).

Results

Participant Enrollment

The sample of 1271 study participants was enrolled from December 2015 to January 2017. The monthly enrollment varied considerably over time, with peak months occurring in the summer and monthly enrollment descriptive statistics as follows: mean 104.8, SD 59.9, median 83, minimum 1, and maximum 179.

Participant Baseline Characteristics

The participant characteristics at baseline are described in Table 4. Consistent with the pattern reported in other studies of internet smoking cessation interventions [10,11], our sample was predominantly female (78%, 991/1271) and aged approximately 45 years. Most participants were married or had a long-term partner (68.3%, 867/1271), had made a quit attempt in last 12 months (75.77%, 963/1271), and had at least a high school degree (72.15%, 917/1271). The only significant between-condition difference in baseline participant characteristics was the larger proportion of participants in the QuitOnline condition who reported having a household member who smoked (37.4%, 235/638) than the MobileQuit condition (31.1%, 195/633).

Participant Internet Usage

The screening procedure validated that all participants had functional smartphone and email service and they actively used both a smartphone and a desktop computer or tablet. A baseline question asked how participants accessed the internet. The results indicated that 56.57% (719/1271) mostly used a smartphone, 13.14% (167/1271) mostly used some other device, 22.42% (285/1271) used both a smartphone and other device equally, 7.71% (98/1271) indicated that it depends, and 0.16% (2/1271) did not know. No between-condition differences were found on these measures.

XSL•FO
RenderX

Table 4. Participant characteristics at screening/baseline by condition.

Participant characteristic[a]	QuitOnline (n=638)	MobileQuit (n=633)	Total (n=1271)
Age (years), mean (SD)	45.6 (12.3)	44.2 (12.9)	44.9 (12.7)
Female, n (%)	500 (78.5)	491 (77.6)	991 (78.0)
Married or have a long-term partner, n (%)	432 (67.7)	435 (68.7)	867 (68.3)
Race/ethnicity, n (%)			
White, non-Hispanic	485 (76.3)	485 (76.9)	970 (76.6)
Other	151 (23.7)	146 (23.1)	297 (23.4)
Education, n (%)			
Not high school graduate	186 (29.2)	168 (26.5)	354 (27.9)
High school graduate/some college	320 (50.2)	337 (53.2)	657 (51.7)
Associate or bachelor's degree	126 (19.7)	125 (19.7)	251 (19.7)
Master's or doctorate degree	6 (0.9)	3 (0.5)	9 (0.7)
Regularly smoked for 4 or more years, n (%)	609 (95.5)	591 (93.4)	1200 (94.4)
Cigarettes/day (previous 6 months), mean (SD)	17.9 (9.9)	17.1 (7.9)	17.5 (8.4)
Quit attempt in last 12 months, n (%)	480 (75.2)	483 (76.3)	963 (75.8)
Currently use other nicotine products, n (%)			
Electronic cigarettes	140 (21.9)	128 (20.2)	268 (21.1)
Cigar	40 (6.3)	45 (7.1)	85 (6.7)
Pipe	4 (0.6)	9 (1.4)	13 (1.0)
Chew/snuff	11 (1.7)	14 (2.2)	25 (2.0)
None	311 (48.7)	305 (48.2)	616 (48.5)
Use quit aids, n (%)			
Nicotine replacement	100 (15.7)	99 (15.6)	199 (15.7)
Prescription medication	41 (6.4)	47 (7.4)	88 (6.9)
Formal treatment	19 (3.0)	9 (1.4)	28 (2.2)
No use	484 (75.9)	493 (77.9)	977 (76.9)
Nicotine dependence (Fagerstrom Test for Nicotine Dependence-6), mean (SD)	5.4 (2.2)	5.5 (2.2)	5.5 (2.2)
Self-efficacy/confidence, mean (SD)	2.5 (1.2)	2.6 (1.1)	2.6 (1.2)
Readiness to quit, mean (SD)	8.8 (1.6)	8.9 (1.6)	8.9 (1.6)
Depression status (Patient Health Questionnaire-8), mean (SD)	9.1 (6.3)	9.2 (6.0)	9.2 (6.1)
Heavy alcohol use, n (%)[b]	58 (9.1)	65 (10.3)	123 (9.7)
Cannabis use in last 7 days, n (%)	98 (15.4)	96 (15.2)	194 (15.3)
Spouse/partner currently smokes, n (%)[c]	189 (30.3)	177 (28.1)	366 (28.8)
Household member currently smokes, n (%)	235 (37.4)	195 (31.1)	430 (34.2)
Number of 5 best friends who smoke, mean (SD)	1.9 (1.6)	1.9 (1.6)	1.9 (1.6)

[a]Participants could refuse to answer any question.

[b]Defined as greater than or equal to 13 drinks/week for men and greater than or equal to 7 drinks/week for women.

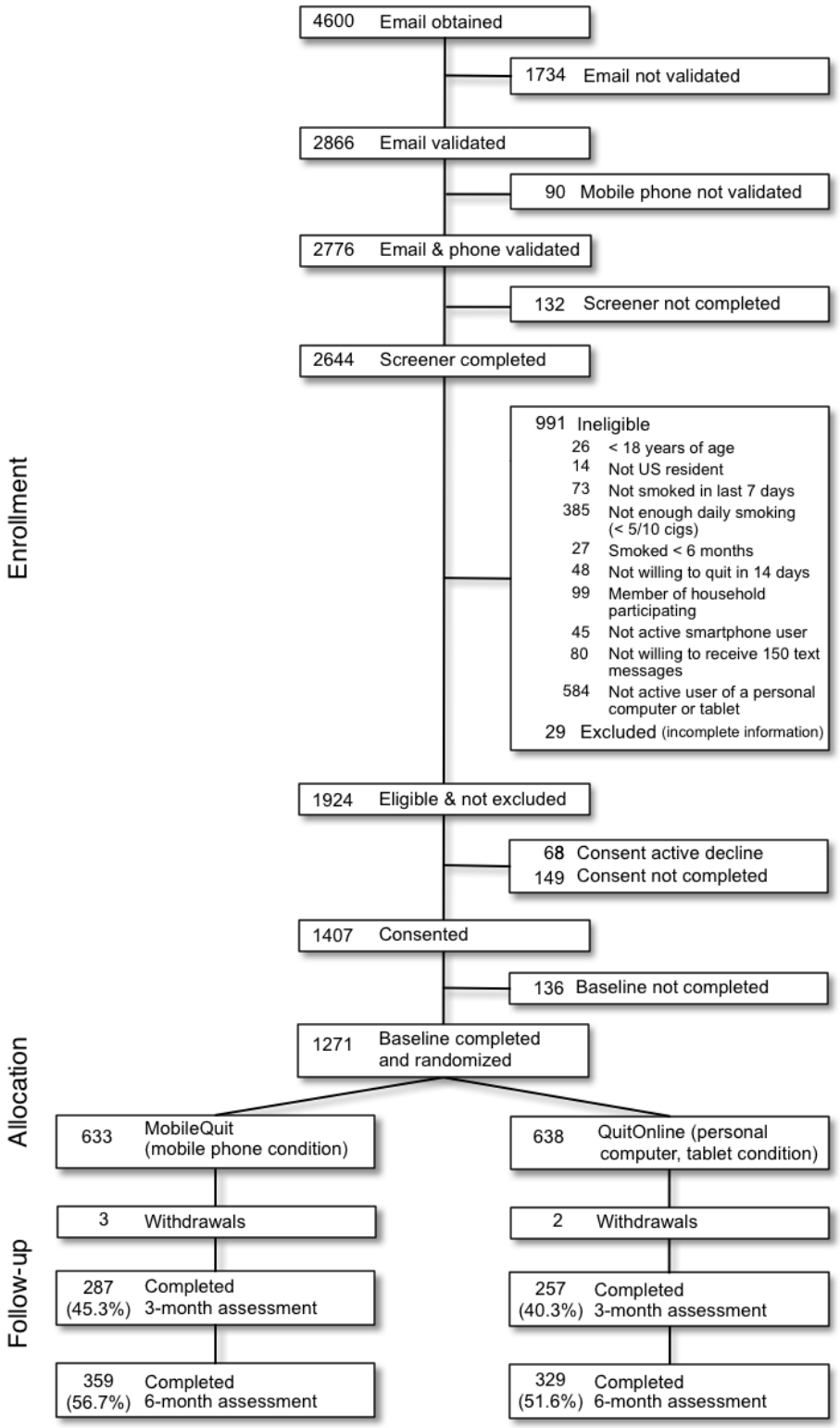
[c]Denominator is full sample, participants without a spouse or with spouses who do not smoke=0 and participants with a spouse who smokes=1.

Participant Flow Through the Study

As shown in the Consolidated Standards of Reporting Trials [48] diagram (Figure 7), of the 1271 study participants initially enrolled, 42.80% (544/1271) completed the 3-month follow-up assessment, 54.13% (688/1271) completed the 6-month follow-up assessment, and 36.43% (463/1271) of participants across conditions completed both assessments.

XSL•FO
RenderX

Figure 7. Consolidated Standards of Reporting Trials diagram depicting flow of participants through the study.



Analysis of baseline characteristics of participants who completed assessments (Complete Cases) failed to reveal any significant differences between conditions. However, the 3-month follow-up assessment was more likely to be completed by participants who were female (46.1% [457/991] compared with 30.8% [86/279]), who reported that they did not have a long-term partner (50.6% [204/403] compared with 39.2% [340/867]), and who reported using a nicotine replacement aid (50.8% [101/199] compared with 41.32% [443/1072]). The 6-month follow-up assessment was more likely to be completed by participants who were older (mean 45.7 years, SD 12.5 compared with mean 43.9 years, SD 12.8), less depressed (PHQ-8 score: mean 8.6, SD 5.8 compared with mean 9.9, SD 6.5), had at least a college degree (61.5% [160/260] compared with 52.2% [528/1011]), used a nicotine replacement aid (61.3% [122/199] compared with 52.8% [566/1072]), and did not have a long-term partner (60% [242/403] compared with 51.4%

XSL•FO
RenderX

[446/867])—especially a partner who smoked (56.1% [498/887] compared with 49.7% [182/366]).

Tobacco Outcomes

The ITT 7-day point prevalence smoking abstinence results across both conditions were 16.05% (204/1271) at 3 months, 21.95% (279/1271) at 6 months, and 12.27% (156/1271) considering both 3 and 6 months (see Table 5). Participants in the MobileQuit condition displayed significantly greater smoking abstinence than those in QuitOnline at 3 months (adjusted OR 2.02, 95% CI 1.48-2.76; *P*<.001), at 6 months (adjusted OR 1.38, 95% CI 1.05-1.80, *P*=.02), and using repeated point prevalence at 3 and 6 months (adjusted OR 1.97, 95% CI 1.39-2.80; *P*<.001).

Complete Case smoking abstinence results across both conditions were 37.5% (204/544) at 3 months, 40.6% (279/688) at 6 months, and 33.7% (156/463) repeated point prevalence abstinence at both 3 and 6 months. MobileQuit participants displayed significantly greater smoking abstinence at 3 months (adjusted OR 2.12, 95% CI 1.48-3.03; *P*<.001) and at both 3 and 6 months (adjusted OR 1.95, 95% CI 1.31-2.91; *P*<.001) but not at 6 months (adjusted OR 1.27, 95% CI 0.93-1.73; *P*=.13). For participants who did not achieve abstinence, changes in the number of cigarettes smoked and number of quit attempts were not detected by condition at the 3- or 6-month follow-up.

Table 5. Smoking abstinence at follow-up assessments by condition.

Type of analysis	3-month assessment	6-month assessment	3- and 6-month assessments
All participants (intention-to-treat)			
MobileQuit (n=633), n (%)	131 (20.7)	156 (24.6)	100 (15.8)
QuitOnline (n=638), n (%)	73 (11.4)	123 (19.3)	56 (8.8)
Between group difference			
Adjusted OR[a] (95% CI)	2.02 (1.48-2.76)	1.38 (1.05-1.80)	1.97 (1.39-2.80)
P value	<.001	0.02	<.001
Participants who completed assessments (Complete Case)			
MobileQuit, n/N (%)	131/287 (45.6)	156/359 (43.5)	100/247 (40.5)
QuitOnline, n/N (%)	73/257 (28.4)	123/329 (34.4)	56/216 (25.9)
Between group difference			
Adjusted OR (95% CI)	2.12 (1.48-3.03)	1.27 (0.93-1.73)	1.95 (1.31-2.91)
P value	<.001	0.128	<.001

[a]OR: odds ratio.

Predictors and Moderators of Tobacco Outcomes

Analyses of baseline sample characteristics as possible predictors of repeated point prevalence abstinence revealed that repeated point abstinence was more likely to be reported by those who have higher levels of self-efficacy (self-confidence; beta=.35; *P*<.001; OR 1.421, 95% CI 1.176-1.718) and less likely for those with friends who smoke (beta=−.14; *P*=.030; OR 0.868, 95% CI 0.763-0.986). Only self-efficacy was significantly associated with repeated point abstinence (beta=.33; *P*=.001; OR 1.386, 95% CI 1.145-1.677) in the multivariate model. No significant interactions between intervention condition and the predictor variables were found.

Text Message Delivery

The participants assigned to MobileQuit were sent a considerable number of text messages (mean 278.51 texts, median 295, SD 71.90, range 6-452). Sending fewer than 200 text messages was associated with 11.1% (70/633) of participants who opted out of receiving messages or who withdrew from the study.

Participant Engagement (Use of Assigned Treatment)

The engagement metrics for all participants are presented in Table 6. The MobileQuit participants (n=633) visited their Web app program an average of 5 times more frequently than did QuitOnline participants (n=638): *z*=−20.33; *P*<.001. Among the MobileQuit participants, 90.0% (570/633) visited multiple times, 6.0% (38/633) visited once, and 4.0% (25/633) never visited. Among the QuitOnline participants, 39.0% (249/638) visited multiple times, 32.0% (204/638) visited once, and 29.0 (185/638) never visited. A different pattern emerged regarding visit duration. Owing to the brief amount of content on MobileQuit pages, 50% of visits to that program lasted ≤25 seconds. As a result, the QuitOnline participants spent significantly more time visiting their program website (*z*=−5.44; *P*<.001).

XSL•FO
RenderX

Table 6. Program visit engagement by condition for all participants (N=1271).

Type of analysis	Mean (SD)	Median
Overall number of program visits		
QuitOnline program visits	2.32 (4.44)	1
MobileQuit program visits	15.92[a] (15.79)	11
Overall duration of program visits (min)		
QuitOnline program visits	21.90 (35.42)	11
MobileQuit program visits	22.34[b] (30.46)	11

[a]Difference in overall number of website visits between QuitOnline and MobileQuit: *P*=.001 (nonparametric Mann-Whitney *U*).

[b]Difference in overall duration of website visits between QuitOnline and MobileQuit: *P*<.001 (nonparametric Mann-Whitney *U*).

Table 7. Visits to Web program by device type and condition.

Type of analysis	QuitOnline visits (n=438), n (%)	MobileQuit visits (n=604), n (%)[a]
Device used for visit		
Nonmobile devices		
Desktop computer	*500 (34.5)* [b]	25 (0.3)
Tablet	*191 (13.2)*	157 (1.6)
Other nonmobile	0 (0)	856 (8.7)
Mobile devices		
Smartphone	607 (41.9)	*7888 (80.0)*
Feature phone	149 (10.3)	*932 (9.5)*
Other mobile device[c]	1 (0.1)	*1 (0)*
Total devices	1448 (100)	9859 (100)
Device recommended or not		
Recommended	*691 (47.7)*	*8821 (89.5)*
Not recommended	757 (523)	1038 (10.5)
Total devices	1448 (100)	9859 (100)

[a]Among the original total of 10,081 MobileQuit visits a device could not be measured in 38 instances and another 184 very short visits were associated with a robot device. The remaining 9859 sessions described in this table represent 97.8% of the original total of MobileQuit visits and 100% of QuitOnline visits.

[b]Text formatted in italics indicate devices classified as fitting the more broadly defined recommended group of devices for each treatment condition (mobile vs nonmobile).

[c]Two visits were recorded—one for each condition—as having been made by a mobile device without any additional details. We listed these 2 episodes in order to provide as comprehensive an account as possible.

The MobileQuit participants were instructed to use a smartphone to visit their program whereas the QuitOnline participants were told to use a desktop computer or tablet. Table 7 describes the devices participants used to visit their program according to the ScientiaMobile WURFL validation tool [44,45], grouped as mobile or nonmobile. Consistent with the MobileQuit's built-in constraint, 89.45% (8820/9859) of the MobileQuit visits were made using the intended mobile device (80% [7888/9859] used a smartphone and 9.45% [932/9859] used a feature phone) whereas 47.7% (691/1448) of QuitOnline visits used the intended nonmobile device (χ_1^2=1645.9; *P*<.001). Analyses of within-participant usage patterns revealed that among the MobileQuit participants, 76.0% (459/604) used only an intended mobile device (primarily a smartphone) across all visits, 23% (139/604) used both mobile and nonmobile devices, and 0.1% (6/604) used only a nonmobile device. Among the QuitOnline participants, 31.3% (137/438) used only an intended nonmobile device across all visits, 16.7% (73/438) used both mobile and nonmobile devices, and 52.1% (228/438) used only a mobile device (primarily a smartphone).

Usability, Helpfulness, and Satisfaction

At 3 months, both programs were described as being easy to use: MobileQuit participants (n=283, mean 3.27, SD 1.04; Somewhat easy=13.1%, Moderately easy=23.0%, and Extremely easy=57.2%) and QuitOnline participants (n=235, mean 3.03, SD 1.26; Somewhat easy=15.3%, Moderately easy=20.4%, and Extremely easy=51.5%). Similar results were obtained for program helpfulness: MobileQuit participants (n=281, mean

XSL•FO
RenderX

2.82, SD 1.20; Somewhat helpful=19.2%, Moderately helpful=27.4%, and Extremely helpful=37.7%); QuitOnline participants (n=234, mean 2.62, SD 1.36; Somewhat helpful=20.1%, Moderately helpful=23.9%, and Extremely helpful=35.5%).

The MobileQuit participants reported that they read 84.5% of the text messages they received (n=278, SD 24.6), that they were satisfied with the number of texts received (n=281, mean 2.46, SD 1.36; Not enough=5%, Just the right number=56.6%, Too many=33.1%, and No opinion=5.0%), and that receiving text messages made it easier for them to quit smoking (n=281, mean 2.46, SD 1.36; Somewhat=18.9%, Moderately=26.0%, and Extremely=29.2%). Significantly more MobileQuit participants (n=286; Yes=87%, No=5%, and Not sure=7%) reported they would recommend their program to "friends or family members who are interested in quitting smoking" than QuitOnline participants (n=253; Yes=80%, No=6%, and Not sure=14%): χ^2=6.6, *P*=.036.

Use of Other Tobacco Products, Quit Aids, and Nonassigned Treatments

The participants in the 2 conditions reported very similar patterns of using other tobacco products, quit aids, and nonassigned quit smoking treatments. The most frequently listed other tobacco products at the 3- and 6-month assessments were *Ecigs* and Other. Among the MobileQuit participants, 15% reported using *Ecigs* at 3 months and 13% at 6 months; 16% reported using *Other* at 3 months and 12% at 6 months. Among the QuitOnline participants, 16% reported using *Ecigs* at 3 months and 12% at 6 months; 19% reported using *Other* at 3 months and 19% at 6 months.

The most frequently listed quit aid at the 3- and 6-month assessments was NRT. Among MobileQuit participants, NRT use was reported by 27% of the participants at 3 months and 19% at 6 months. Prescription use was reported by 9% of the participants at 3 months and 6% at 6 months, and Formal quit smoking treatment use was reported by 4% of the participants at 3 months and 3% at 6 months. Among the QuitOnline participants, NRT use was reported by 21% at 3 months and 17% at 6 months. Prescription use was reported by 10% of the participants at 3 months and 6% at 6 months, and Formal quit smoking treatment use was reported by 5% at 3 months and 3% at 6 months.

Discussion

Principal Findings

Overall, the smoking cessation rates and absolute smoking abstinence levels for the 2 well-matched (fully automated, best practice content) smoking cessation programs were consistent with results reported in meta-analyses of other internet smoking cessation interventions [10,11,16,49]. However, the MobileQuit intervention for mobile devices was significantly more effective in encouraging smoking cessation than the QuitOnline designed for use on devices other than mobile devices. Specifically, ITT results of the MobileQuit participants displayed significantly greater smoking cessation than the QuitOnline participants at all follow-up assessments: 20.7 vs 11.4% at 3 months, 24.6% vs 19.3% at 6 months, and 15.8% vs 8.8% at 3 and 6 months. Similarly, Complete Case results significantly favored MobileQuit at both 3 months (45.6% vs 28.4%) and the combined 3- and 6-month assessments (40.5% vs 25.9%). However, the advantage for MobileQuit (43.5% vs 34.4%) at 6 months did not reach significance.

The participants in each condition found their treatment program acceptable, both in terms of helpfulness (MobileQuit=84.3%; QuitOnline=79.5%) and ease of use (MobileQuit=87.2%; QuitOnline=93.3%). The small between-condition differences in helpfulness and ease of use did not reach significance. Significantly more participants in MobileQuit than in QuitOnline (87.4% vs 79.8%) reported they would recommend the program to their friends/family interested in quitting.

There were 2 striking differences in usage pattern between the 2 intervention groups. First, not surprisingly, because the intervention included a built-in validation tool designed to try to constrain its use to smartphones, almost all MobileQuit visits occurred using that recommended device. In marked contrast, visits to the QuitOnline program—which did not constrain device type—showed considerable variability in being accessed using recommended as well as nonrecommended devices (including smartphones and other mobile devices).

Second, the MobileQuit participants visited their program website an average of 5 times more often than the QuitOnline participants. Stated differently, the MobileQuit intervention was used more frequently but in smaller doses/shorter visits compared with the QuitOnline intervention.

Limitations

There are several limitations of this study that are worth noting. First, self-reported smoking abstinence was not validated by biochemical measures. However, most published tobacco cessation programs rely on self-reported data, and Glasgow et al [50] as well as the Society for Research on Nicotine and Tobacco (SRNT) Subcommittee on Biochemical Verification [51] has recommended that biochemical validation need not be required when a study's self-help design makes it impractical, when demand characteristics are not likely to differentially affect reports by condition, or when accurate estimates of tobacco use can be obtained using multiple self-reported measures.

The participants in this study were an average age of 45 years and 78% were female. Although this profile is similar to participant characteristics reported in a number of other internet-based smoking cessation studies [5,10,52,53], our observed results may not generalize to younger smokers. As participants in this study agreed to be assigned to either of the internet-based smoking cessation interventions, it is also possible that our study results may not generalize to smokers who would have preferred to use only a smartphone-delivered intervention or, alternatively, only an internet intervention that did not use a smartphone. The study design that yoked users to particular types of computer devices (eg, smartphones for MobileQuit participants) may have been less convenient and possibly resulted in lower engagement and efficacy, compared with using responsive design technology that would have enabled either

XSL•FO
RenderX

of the interventions to be used interchangeably on any internet-accessible computer device (desktop, laptop, tablet, and smartphone) [23].

There was also substantial assessment attrition at follow-up which, although consistent with results reported by other smoking cessation studies [10], was somewhat greater than has been reported for mobile smoking cessation interventions [16]. We also observed that 29% of the QuitOnline participants never visited their assigned program (an extreme case of nonusage attrition, [54]) compared with only 4% of the MobileQuit participants who did not visit. It is also helpful to consider these findings from the perspective of other published results. For example, QuitOnline's 29% nonvisit rate is similar to the 20% to 25% nonvisit rate results reported by Cobb and Graham [55] for 2005 participants in a smoking cessation randomized controlled trial (RCT). This difference is also consistent with our expectation that the push or proactive outreach of MobileQuit's text messages would encourage relatively more participant engagement. Moreover, by providing time-limited access to its Major Topics content, the MobileQuit program may have increased its perceived value of the program (the scarcity principle) to its users. The program's tunnel design also paced program content so that it was better synchronized with the phases of quitting, which may have encouraged involvement and sustained interest.

Strengths

This trial is one of the relatively few large-scale (N=1271) RCTs examining the efficacy of smartphone-delivered smoking cessation intervention. The interventions were designed to reliably and unobtrusively track the device used to access the program as well as the frequency and duration of each participant's contact. Thus, the trial represents a rare instance of both describing the devices that participants used to access the internet intervention and comparing the usage pattern across device type. The use of a browser-based Web app for MobileQuit enabled us to avoid having to create 2 altogether different native apps (one for iOS devices [iPhones and iPads] and another for Android devices [smartphones and tablets]), to avoid the review and delay associated with distributing native apps via official company app stores [23], and to set the stage for a responsive design that would be usable across all devices and their operating systems. These benefits combined with the emergence of the more sophisticated *progressive Web app* [56] are encouraging more widespread use of the Web app design approach. As text messages are increasingly being delivered on nonmobile devices, a Web app plus text messaging intervention can be delivered across all devices.

Future Directions

Although the absolute proportion of smokers who quit in this study was encouraging, additional research is warranted that examines how to encourage even more widespread smoking abstinence. This design did not permit a direct analysis of the individual and combined features of the 2 conditions. For example, impact of device type and use of text messaging on long-term smoking abstinence could be examined using a completely crossed 2×2 design (smartphone vs not smartphone and app + text messaging vs app only). It would also be helpful to examine the likely contributions of other factors that were not examined directly in our RCT (eg, tunnel vs hybrid matrix-hierarchical information architecture). In all instances, research should assess the devices and the usage patterns that participants follow to access their internet interventions.

Conclusions

Despite the fact that this study did not pinpoint the exact design feature(s) that explain the increased efficacy of MobileQuit over QuitOnline, the study nonetheless provides evidence for the benefit of optimizing an intervention design for smartphones and other mobile devices over a usual care internet intervention designed primarily for use on nonmobile devices such as desktops, laptops, or tablets. Our study also helps to underscore that participants will use multiple devices (what Google describes as *sequential screening* [22]) irrespective of recommendations to do otherwise. As a result, we assert that future internet interventions should be designed for use on all of the devices that users prefer to access the internet. Essentially, in the increasingly multiscreen world, the approach of designing internet interventions for mobile vs nonmobile devices will be replaced by responsive-designed programs that share user data across devices and embody pervasive information architecture [22,23,57,58].

Acknowledgments

The authors wish to express their appreciation to Edward Lichtenstein, PhD, and Herbert H Severson, PhD, (both at Oregon Research Institute) for their wise and generous counsel and to Tim Woolley and his colleagues at IEQ Technology (Springfield, OR) for assistance in developing the MobileQuit and QuitOnline interventions and their related research components. The authors also thank other members of their ORI research team: Katie Clawson and Margaret Ortiz. Finally, the authors acknowledge the contribution of Kevin Danaher who designed and managed their successful internet-based participant enrollment marketing campaign.

This work was supported by a research grant R01CA172205 from the National Cancer Institute (US National Institutes of Health).

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Multistep enrollment validation sequence for the MobileQuit randomized controlled trial.

[PNG File, 118KB - jmir_v21i6e13290_app1.png]

Multimedia Appendix 2

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 590KB - jmir_v21i6e13290_app2.pdf]

References

1. Brendryen H, Drozd F, Kraft P. A digital smoking cessation program delivered through internet and cell phone without nicotine replacement (happy ending): randomized controlled trial. J Med Internet Res 2008;10(5):e51 [FREE Full text] [doi: 10.2196/jmir.1005] [Medline: 19087949]
2. Brendryen H, Kraft P. Happy Ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-84; discussion 485. [doi: 10.1111/j.1360-0443.2007.02119.x] [Medline: 18269367]
3. Etter J. Comparing the efficacy of two internet-based, computer-tailored smoking cessation programs: a randomized trial. J Med Internet Res 2005;7(1):e2 [FREE Full text] [doi: 10.2196/jmir.7.1.e2] [Medline: 15829474]
4. Japuntich SJ, Zehner ME, Smith SS, Jorenby DE, Valdez JA, Fiore MC, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine Tob Res 2006 Dec;8(Suppl 1):S59-S67. [Medline: 17491172]
5. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. J Med Internet Res 2008;10(5):e40 [FREE Full text] [doi: 10.2196/jmir.993] [Medline: 19017582]
6. Strecher VJ, Shiffman S, West R. Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. Addiction 2005 May;100(5):682-688. [doi: 10.1111/j.1360-0443.2005.01093.x] [Medline: 15847626]
7. Swartz LH, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [doi: 10.1136/tc.2003.006189] [Medline: 16436397]
8. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
9. 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]
10. Taylor GMJ, Dalili MN, Semwal M, Civljak M, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2017 Sep 4;9:CD007078. [doi: 10.1002/14651858.CD007078.pub5] [Medline: 28869775]
11. Graham AL, Carpenter KM, Cha S, Cole S, Jacobs MA, Raskob M, et al. Systematic review and meta-analysis of internet interventions for smoking cessation among adults. Subst Abuse Rehabil 2016;7:55-69 [FREE Full text] [doi: 10.2147/SAR.S101660] [Medline: 27274333]
12. Intille SS, Kukla C, Farzanfar R, Bakr W. Just-in-time technology to encourage incremental, dietary behavior change. AMIA Annu Symp Proc 2003:874 [FREE Full text] [Medline: 14728379]
13. 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]
14. Free C, Whittaker R, Knight R, Abramsky T, Rodgers A, Roberts IG. Txt2stop: a pilot randomised controlled trial of mobile phone-based smoking cessation support. Tob Control 2009 Apr;18(2):88-91. [doi: 10.1136/tc.2008.026146] [Medline: 19318534]
15. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin R, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [doi: 10.1136/tc.2005.011577] [Medline: 16046689]
16. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
17. Graham AL, Jacobs MA, Cohn AM, Cha S, Abroms LC, Papandonatos GD, et al. Optimising text messaging to improve adherence to web-based smoking cessation treatment: a randomised control trial protocol. BMJ Open 2016 Mar 30;6(3):e010687 [FREE Full text] [doi: 10.1136/bmjopen-2015-010687] [Medline: 27029775]
18. Pfaeffli DL, Dobson R, Whittaker R, Maddison R. The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: a systematic review. Eur J Prev Cardiol 2016 May;23(8):801-817. [doi: 10.1177/2047487315613462] [Medline: 26490093]

XSL•FO
RenderX

19. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Jun;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]
20. 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]
21. Pew RC. Pew Research Center. 2018. Internet/Broadband fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 77rAFSY7m]
22. Google. 2012. The new multi-screen world: Understanding cross-platform consumer behavior URL: http://services.google.com/fh/files/misc/multiscreenworld_final.pdf
23. Danaher BG, Brendryen H, Seeley JR, Tyler MS, Woolley T. From black box to toolbox: outlining device functionality, engagement activities, and the pervasive information architecture of mHealth interventions. Internet Interv 2015 Mar 1;2(1):91-101. [doi: 10.1016/j.invent.2015.01.002] [Medline: 25750862]
24. Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [doi: 10.2196/jmir.7.2.e12] [Medline: 15914459]
25. Pugatch J, Grenen E, Surla S, Schwarz M, Cole-Lewis H. Information architecture of web-based interventions to improve health outcomes: systematic review. J Med Internet Res 2018 Mar 21;20(3):e97 [FREE Full text] [doi: 10.2196/jmir.7867] [Medline: 29563076]
26. Danaher BG, Severson HH, Crowley R, van Meter N, Tyler MS, Widdop C, et al. Randomized controlled trial examining the adjunctive use of nicotine lozenges with MyLastDip: an eHealth smokeless tobacco cessation intervention. Internet Interv 2015 Mar;2(1):69-76. [doi: 10.1016/j.invent.2014.12.004]
27. Danaher BG, Severson HH, Andrews JA, Tyler MS, Lichtenstein E, Woolley TG, et al. Randomized controlled trial of MyLastDip: a Web-based smokeless tobacco cessation program for chewers ages 14-25. Nicotine Tob Res 2013 Sep;15(9):1502-1510 [FREE Full text] [doi: 10.1093/ntr/ntt006] [Medline: 23410803]
28. Nielsen J. Nielsen Norman Group. How many test users in a usability study? URL: https://www.nngroup.com/articles/how-many-test-users/ [accessed 2019-01-02] [WebCite Cache ID 758PST7if]
29. Nielsen J. Nielsen Norman Group. 2012. Thinking aloud: The #1 usability tool URL: https://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/ [accessed 2019-01-02] [WebCite Cache ID 758POiIOV]
30. Nielsen Norman Group. 2014. Turn user goals into task scenarios for usability testing URL: https://www.nngroup.com/articles/task-scenarios-usability-testing/ [accessed 2019-01-02] [WebCite Cache ID 758PV39DY]
31. Usability.gov. Use case scenarios URL: https://www.usability.gov/how-to-and-tools/methods/use-cases.html [accessed 2019-01-02] [WebCite Cache ID 758PY6Fku]
32. Lew GS. What do users really do? Experience sampling in the 21st century. In: Jacko JA, editor. Human-Computer Interaction. New Trends: 13th International Conference, HCI International 2009, San Diego, CA, USA, July 19-24, 2009, Proceedings, Part I (Lecture Notes in Computer Science). Berlin: Springer; 2009.
33. Lew G. User Experience Magazine. The truth is out there: using mobile technology for experience sampling URL: http://uxpamagazine.org/mobile_technology_experience_sampling/ [accessed 2019-01-02] [WebCite Cache ID 758Pe0UVb]
34. Fagerstrom KO, Schneider NG. Measuring nicotine dependence: a review of the Fagerstrom Tolerance Questionnaire. J Behav Med 1989 Apr;12(2):159-182. [Medline: 2668531]
35. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
36. Transdisciplinary Tobacco Use Research Center (TTURC) Tobacco Dependence, Baker TB, Piper ME, McCarthy DE, Bolt DM, Smith SS, et al. Time to first cigarette in the morning as an index of ability to quit smoking: implications for nicotine dependence. Nicotine Tob Res 2007 Nov;9(Suppl 4):S555-S570 [FREE Full text] [doi: 10.1080/14622200701673480] [Medline: 18067032]
37. Danaher BG, Smolkowski K, Seeley JR, Severson HH. Mediators of a successful web-based smokeless tobacco cessation program. Addiction 2008 Oct;103(10):1706-1712 [FREE Full text] [doi: 10.1111/j.1360-0443.2008.02295.x] [Medline: 18715238]
38. Biener L, Abrams DB. The Contemplation Ladder: validation of a measure of readiness to consider smoking cessation. Health Psychol 1991;10(5):360-365. [Medline: 1935872]
39. 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]
40. Wells TS, Horton JL, LeardMann CA, Jacobson IG, Boyko EJ. A comparison of the PRIME-MD PHQ-9 and PHQ-8 in a large military prospective study, the Millennium Cohort Study. J Affect Disord 2013 May 15;148(1):77-83. [doi: 10.1016/j.jad.2012.11.052] [Medline: 23246365]
41. Kroenke K, Strine TW, Spitzer RL, Williams JB, 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]
42. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatric Annals 2002 Sep 1;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]

XSL•FO
RenderX

43. Danaher BG, Boles SM, Akers L, Gordon JS, Severson HH. Defining participant exposure measures in Web-based health behavior change programs. J Med Internet Res 2006;8(3):e15 [FREE Full text] [doi: 10.2196/jmir.8.3.e15] [Medline: 16954125]
44. ScientiaMobile. 2019. URL: https://www.scientiamobile.com/ [accessed 2019-04-23] [WebCite Cache ID 77r9RfEI2]
45. ScientiaMobile. Device Detection URL: https://www.devicedetection.com/ [accessed 2019-04-23] [WebCite Cache ID 77rA4pU1y]
46. Mungovan B. Adobe Blog. Are tablets mobile devices? How will Google's changes in AdWords impact advertisers? URL: https://theblog.adobe.com/are-tablets-mobile-devices-how-will-googles-changes-in-adwords-impact-advertisers/ [accessed 2019-04-23] [WebCite Cache ID 77r9sIGER]
47. Smolkowski K, Danaher BG, Seeley JR, Kosty DB, Severson HH. Modeling missing binary outcome data in a successful web-based smokeless tobacco cessation program. Addiction 2010 Jun;105(6):1005-1015 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02896.x] [Medline: 20148782]
48. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. PLoS Med 2010 Mar 24;7(3):e1000251 [FREE Full text] [doi: 10.1371/journal.pmed.1000251] [Medline: 20352064]
49. McCrabb S, Baker AL, Attia J, Skelton E, Twyman L, Palazzi K, et al. Internet-based programs incorporating behavior change techniques are associated with increased smoking cessation in the general population: a systematic review and meta-analysis. Ann Behav Med 2018 May 10:180-195. [doi: 10.1093/abm/kay026] [Medline: 29750240]
50. Glasgow RE, Mullooly JP, Vogt TM, Stevens VJ, Lichtenstein E, Hollis JF, et al. Biochemical validation of smoking status: pros, cons, and data from four low-intensity intervention trials. Addict Behav 1993;18(5):511-527. [Medline: 8310871]
51. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
52. Cobb NK, Jacobs MA, Wileyto P, Valente T, Graham AL. Diffusion of an evidence-based smoking cessation intervention through Facebook: a randomized controlled trial. Am J Public Health 2016 Jun;106(6):1130-1135. [doi: 10.2105/AJPH.2016.303106] [Medline: 27077358]
53. Saul JE, Amato MS, Cha S, Graham AL. Engagement and attrition in Internet smoking cessation interventions: Insights from a cross-sectional survey of "one-hit-wonders". Internet Interv 2016 Sep;5:23-29 [FREE Full text] [doi: 10.1016/j.invent.2016.07.001] [Medline: 30135803]
54. 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]
55. Cobb CO, Graham AL. Use of non-assigned interventions in a randomized trial of internet and telephone treatment for smoking cessation. Nicotine Tob Res 2014 Oct;16(10):1289-1297. [doi: 10.1093/ntr/ntu066] [Medline: 24812022]
56. Ater T. Building Progressive Web Apps: Bringing the Power of Native to the Browser. Sebastopol, CA: O'Reilly Media; 2017.
57. Greenfield A. Everyware: The Dawning Age Of Ubiquitous Computing. Berkeley, CA: New Riders Publishing; 2019.
58. Resmini A, Rosati L. Pervasive Information Architecture: Designing Cross-Channel User Experiences. Burlington, MA: Morgan Kaufmann; 2011.

Abbreviations

ITT: intention-to-treat
NRT: nicotine replacement therapy
OR: odds ratio
PHQ-8: Patient Health Questionnaire-8
RCT: randomized controlled trial
WURFL: Wireless Universal Resource FiLe

Edited by G Eysenbach; submitted 03.01.19; peer-reviewed by A Graham, B Haskins, N Nguyen, J Dahne, C Carrion; comments to author 30.01.19; revised version received 24.04.19; accepted 12.05.19; published 06.06.19.

Please cite as:
Danaher BG, Tyler MS, Crowley RC, Brendryen H, Seeley JR
Outcomes and Device Usage for Fully Automated Internet Interventions Designed for a Smartphone or Personal Computer: The MobileQuit Smoking Cessation Randomized Controlled Trial
J Med Internet Res 2019;21(6):e13290
URL: https://www.jmir.org/2019/6/e13290/
doi:10.2196/13290
PMID:31172967

XSL•FO
RenderX

©Brian G Danaher, Milagra S Tyler, Ryann C Crowley, Håvar Brendryen, John R Seeley. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Data Mining in the Development of Mobile Health Apps: Assessing In-App Navigation Through Markov Chain Analysis

Jeroen Stragier[1], PhD; Gilles Vandewiele[2], Eng; Paulien Coppens[3], MSc; Femke Ongenae[2], PhD; Wendy Van den Broeck[3], PhD; Filip De Turck[2], PhD; Lieven De Marez[1], PhD

[1]imec-mict, Department of Communication Sciences, Ghent University, Ghent, Belgium

[2]imec-IDLab, Department of Information Technology, Ghent University, Ghent, Belgium

[3]imec-smit, Department of Communication Sciences, Vrije Universiteit Brussel, Brussels, Belgium

Corresponding Author:
Jeroen Stragier, PhD
imec-mict
Department of Communication Sciences
Ghent University
Miriam Makeplein 1
Ghent, 9000
Belgium
Phone: 32 92649745
Email: jeroen.stragier@ugent.be

Abstract

Background: Mobile apps generate vast amounts of user data. In the mobile health (mHealth) domain, researchers are increasingly discovering the opportunities of log data to assess the usage of their mobile apps. To date, however, the analysis of these data are often limited to descriptive statistics. Using data mining techniques, log data can offer significantly deeper insights.

Objective: The purpose of this study was to assess how Markov Chain and sequence clustering analysis can be used to find meaningful usage patterns of mHealth apps.

Methods: Using the data of a 25-day field trial (n=22) of the *Start2Cycle* app, an app developed to encourage recreational cycling in adults, a transition matrix between the different pages of the app was composed. From this matrix, a Markov Chain was constructed, enabling intuitive user behavior analysis.

Results: Through visual inspection of the transitions, 3 types of app use could be distinguished (route tracking, gamification, and bug reporting). Markov Chain–based sequence clustering was subsequently used to demonstrate how clusters of session types can otherwise be obtained.

Conclusions: Using Markov Chains to assess in-app navigation presents a sound method to evaluate use of mHealth interventions. The insights can be used to evaluate app use and improve user experience.

(J Med Internet Res 2019;21(6):e11934) doi:10.2196/11934

KEYWORDS

eHealth; mHealth; Markov Chain; log data; data analytics

Introduction

Background

The development of wearable technology, particularly in health care, medical, and fitness contexts, provides people with devices that give detailed information about various aspects of their health, for example, by registering users' daily physical activity in terms of step counts, calories, or by providing detailed information about exercise parameters [1].

At the same time, mobile phone technology is reaching significant adoption and penetration rates in both developed and developing countries [2]. This offers an excellent platform to reach a large audience with health-oriented mobile phone apps. Most mobile phones contain sensors that afford measurement of parameters similar to those measured by wearables, including accelerometer and global positioning system.

Mobile health (mHealth) apps are increasingly available to people worldwide. Over 318,000 mHealth apps are estimated

XSL•FO
RenderX

to be available in app stores around the world [3]. These apps target various conditions and behaviors including nutrition, smoking cessation, mental conditions, diabetes, and physical activity [4-6]. However, although mHealth apps can empower people to take their health more into their own hands [7,8], research indicates that the uptake of these apps does not necessarily result in sustained usage [9,10]. In digital behavior change interventions (DBCIs), which use mHealth apps to promote health behavior, this lack of engagement with the technology and intervention too often results in high levels of dropout [11,12].

Engagement with DBCIs can be considered from different perspectives. Perski et al [12] define engagement both in terms of a subjective experience and a behavior. The first considers engagement as the user's subjective experience with the technology, with quality of experience, immersion, usability, and enjoyment as the main determinants of active and frequent usage, whereas the latter takes more objective parameters into account such as frequency and duration of use or use of specific content of the intervention. Yardley et al [11] distinguish between micro and macro levels of engagement, with microengagement referring to specific interaction with the technology and macro levels of engagement considering the engagement with technology within the overall aim of the DBCI to encourage behavior change. This study focuses specifically on this micro level of the engagement with the *Start2Cycle* app, developed to promote recreational cycling. Therefore, it was decided to only analyze the log data of the app and not take any cycling outcome data of the participants into account. Furthermore, as the field trial in which the app was tested was a usability test rather than a full randomized controlled trial, using the outcome data could be misleading in terms of effectiveness.

Measuring engagement with DBCIs receives ample attention. To date however, no consistent or standardized methodology to quantify and measure participant engagement in DBCIs is available [11]. In addition to self-report-based questionnaires distributed after usage of the app, the analysis of log data (anonymous records of real-time actions performed by each user) has been suggested as an opportunity to study actual usage of apps both in the development phase and when the final version has been launched [13] and to tailor interventions [14]. In human-computer interaction research, data mining and machine learning techniques are frequently used for various purposes such as finding user activity patterns in apps and games [15-17], user segmentation [18], and app development [19]. In mHealth research, however, analysis of log data remains often limited to descriptive analysis in terms of number of visits, length of visits, and page views. Although these statistics offer an initial insight into *when* and *how often* participants use mHealth apps, they do not provide us with more profound insights on *how* participants are using them. This information is essential to the development of mHealth interventions and the attainment of their anticipated health effects. Limited research applies more detailed analysis of log data of electronic health apps. Arden-Close [20], for example, used sequence clustering to develop a visualization tool to monitor the use of a Web-based weight management intervention. Furthermore, Miller et al [21] present a framework for analyzing engagement data of health interventions, including suggestions for analyzing log data.

Objectives

This study partly builds on these insights by adding Markov Chain analysis to the toolkit for analyzing log data of mHealth interventions. In multimedia research contexts, Markov Chains have been frequently used to study Web page navigation [22-24]. In these analyses, Markov Chains are used to model a *trail*, that is, an order of Web page views by 1 user within a Web-use session. This study aims to be a methodological example of how this thinking was applied to in-app navigation such that the Markov Chains model the order of page views made by a user within an mHealth app. This provides insight into how the app is used, that is, which types of usage sessions can be identified and if user sessions can be clustered into distinct groups.

Methods

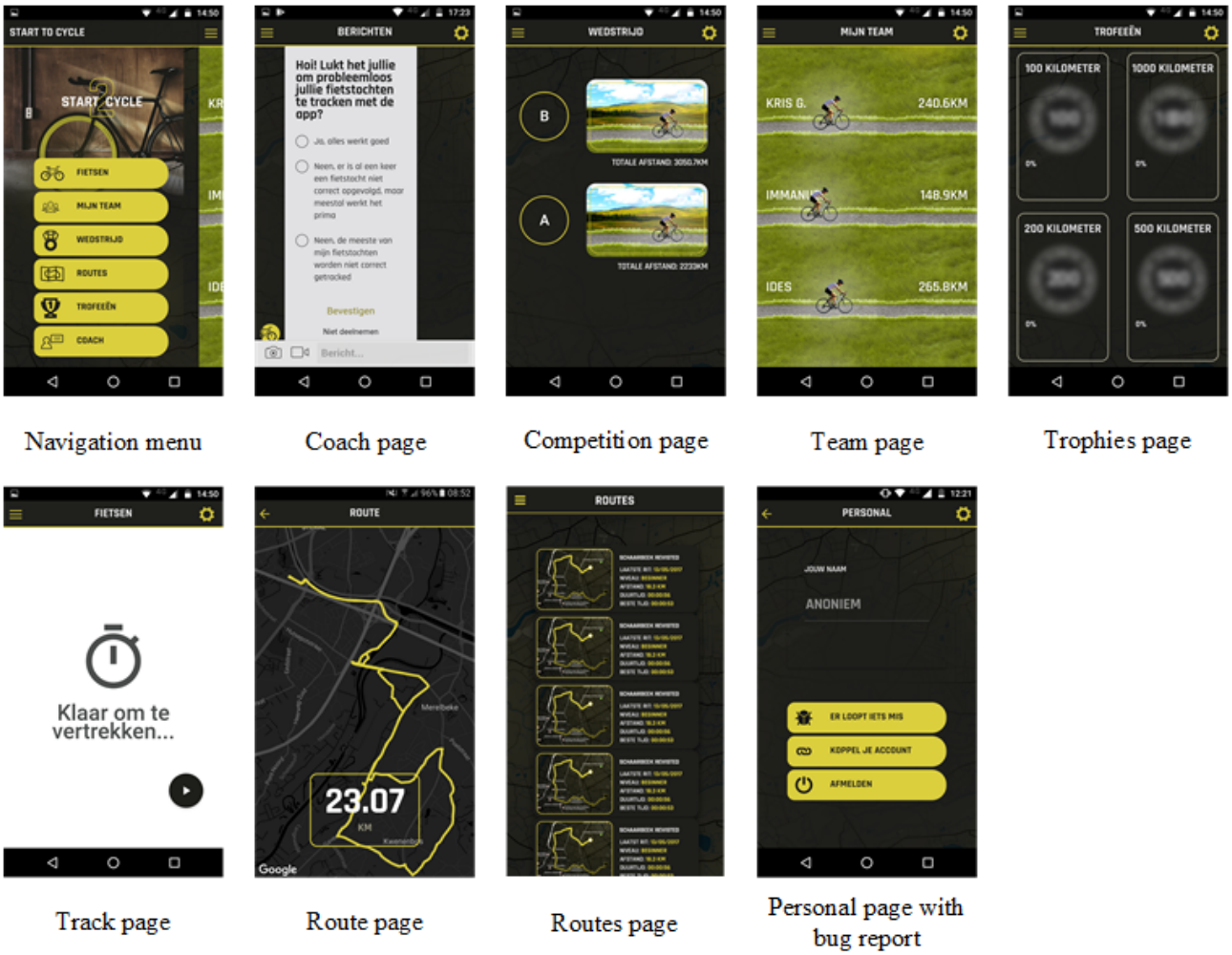
This section details the content of the *Start2Cycle* app, the setup of the field trial in which the app was tested, and the methods applied to the app log data.

The Start2Cycle App

The *Start2Cycle* app was developed to promote recreational cycling in adult populations. The app was designed for both novice and experienced cyclists. It was developed by the Flemish public broadcaster Vlaamse Radio- en Televisieomroeporganisatie, Ghent University, and Vrije Universiteit Brussel and was available for both Android and iPhone operating system devices.

XSL•FO
RenderX

Figure 1. Screenshots of the Start2Cycle app pages.



The app, depicted in Figure 1, has a starting page called the *Coach*. Through this page, 2 researchers provided additional cycling-related content to the app during the field trial by postings polls, news messages, route suggestions, pictures, and movies. Push notifications notified users of new posts. Users could ask questions to the coach. Beyond the starting page, the app had 2 core functionalities: (1) Self-monitoring and feedback on cycling behavior and (2) Rewards and social competition. The latter can be classified as gamification features. The self-monitoring features of the app enable the users to log their cycling activities (*Track* page) and afterwards see the route they took (*Route* page). They were also provided with a page that lists all of the routes they had taken since they started using the app (*Routes* page). Gamification and social features are assumed to have potential to create more engagement with an app and potentially longer lasting behavior change effects [25,26]. Therefore, gamification and social competition and cooperation were included in the app. Users could, for example, collect short videos starring a well-known Flemish sports journalist providing motivational or informational content about cycling on the *Trophies* page. These videos were unlocked at certain *milestones* in the cycling progress of the participant, for example, after his or her first ride and at 25 cycled kilometers. Furthermore, the participants were randomly assigned to 1 of 2 groups (team A and team B), and a competition was set up between the 2 groups. The group with the highest distance cycled (in kilometers) at the end of the trial won the competition. The participants could monitor the performance of their group in comparison to the other on the *Competition* page and the individual performances within their group on the *Team* page of the app. By combining both cooperation among group members and competition between the groups, it was aimed to make the app appealing to both competitive and noncompetitive participants [27,28].

Finally, there was a *Personal* page on which the participants could change their settings, enter or change their screen name, and report errors.

Field Trial

A 25-day field trial of a prototype version of the *Start2Cycle* app, with primarily a usability testing purpose, was set up with 22 participants in Flanders, Belgium in September 2017. The participants were recruited from a group of people that registered for a 1-week cycling holiday that started at the end of the field trial. No inclusion criteria were set. Anyone joining the cycling holiday could participate in the trial. Participants were asked to log every cycling activity during the 25 days preceding the cycling holiday. The project was approved by the institutional review board, and informed consent was obtained from the participants.

Every action the participants performed in the *Start2Cycle* app was logged. More specifically, per user, each action (or *click*)

XSL•FO
RenderX

was logged, along with the time stamp. No session identifier (ID) was collected. Consequently, sessions were identified by grouping actions of a user together, between which no longer than 30 min of inactivity occurred. Using this heuristic, 824 sessions were identified. At the end of the field trial, an online questionnaire was sent to the users to evaluate the use of the app in terms of perceived usefulness, ease of use, motivational potential, enjoyment, and informativeness. The evaluation was conducted using 7-point Likert scales ranging from 1=*Totally disagree* to 7=*Totally agree*.

Markov Chain Analysis

The log data of the *Start2Cycle* app were analyzed by means of Markov Chain analysis. Markov Chains are a sound method to model stochastic processes and have already been successfully adopted in a wide spectrum of applications ranging from auto-completion while typing [25] to the world-renowned ranking algorithm of the Google search engine [29]. They are defined by a state space *S*, a transition matrix *P*, and optionally an initial state distribution π. In the case of this study, we have a discrete state space, defined by the *n* different pages in the mobile app ($S=\{s_1,..., s_n\}$). Moreover, we assume the Markov Chain to be of order *1*. In other words, we assume that the probability to transition to a next state depends only on the current state and not on the history of previous states.

To create a Markov Chain, a transition matrix *P*, in which each element on row *i* and column *j* ($p_{i,j}$) indicates the probability that a user moves from page s_i to page s_j, must be constructed. This transition matrix can be constructed from raw data logs containing a *time stamp*, *session identifier*, and the *performed clicks (source* and *destination)*. The session identifier is optional and can be calculated using a heuristic based on the *user identifier* and time stamp of each data log. The number of transitions $s_i \rightarrow s_j$ can then be counted to form a matrix *M*. Afterwards, the values in *M* are normalized such that each row sums to 1 to result in *P*.

Furthermore, a surrogate exit-state s_{n+1} to which a user transitions when he leaves the app (*exit* page) was created. The initial state distribution π can be estimated by counting the first visited state of every session.

Sequence Clustering

When the number of pages in an app becomes large, manual interpretation of the visualized Markov Chain can become cumbersome. An advantage of having a Markov Chain as underlying data structure of the visualization is that various automated analyses are possible. These can be used to automatically extract insights or to confirm the findings of a researcher. Clustering of different sequences (or trails of user actions) is 1 such analysis [30,31]. To do this, an expectation-maximization algorithm [32] can be applied, which builds a representative Markov Chain per cluster in an iterative fashion until convergence is reached.

Given a collection of sequences *C*, in which each element is represented by a trail of user actions ($\{s_1,..., s_i,..., s_{n+1}\}$, with s_{n+1} the exit-state), we apply the following steps:

1. Determine the number of clusters: *K*
2. Assign each sequence in *C* to a random cluster c_i with $i \in \{1,..., K\}$
3. For each c_i, construct a Markov Chain or transition matrix (P_i) based on its assigned sequences.
4. For each sequence in *C*, calculate the likelihood $L(\square, c_i)$ that the sequence is generated by each Markov Chain by multiplying the probabilities of the different transitions (eg, $L(\{s_0, s_3, s_1, s_{n+1}\}, c=p_{0,3} * p_{3,1} * p_{1,n+1})$. Then assign the sequence to the cluster with the maximal likelihood: $argmax_i(L(\square, c_i))$.
5. Reconstruct each Markov Chain, based on the assigned sequences and repeat steps 4 and 5 until convergence. To check convergence, the assignment of sequences to clusters can be compared with the assignment of the previous iteration.

The output of this algorithm is a mapping from each sequence to a cluster c_i and *K* different Markov Chains that represent each cluster. These Markov Chains can be used to classify new unseen sequences of actions.

Web-Based Visualization Tool

To ensure a more intuitive interpretation of such results for researchers, interactive visualizations of the results can be created. For this study, such an interactive visualization was developed using D3.js [33], a javascript library, and can thus be displayed in a Web browser. The tool offers both static and interactive visualization to the researcher, which afford adaptation of the visualization in terms of (1) a slider that defines a threshold for a minimum value the corresponding probability of an edge must have to be displayed, (2) enabling and disabling certain pages or nodes and corresponding edges from the visualization, and (3) the ability to drag nodes around to create a nice structured topology. Moreover, simulations of user sessions, based on the provided data, can be performed, and the results of the sequence clustering can be inspected.

Results

This section details the results of the analysis. First, an overview of the sample is given followed by the results of the Markov Chain and sequence clustering analyses.

Descriptive Statistics

The app was tested by 22 participants, 18 of which (82%) were male and 12 of which (55%) had a higher education. The evaluation survey, distributed at the end of the field trial, was completed by 16 users (73%). The mean age was 50 years (SD 11.17). Survey results indicated that the app was positively evaluated by the participants. It was considered to be fairly useful (mean 4.63, SD 1.15), easy to use (mean 4.50, SD 1.46), and fun (mean 4.38, SD 1.31). Users rated the app as somewhat motivating (mean 4.25, SD 1.39). The app was scored lower in terms of informativeness (mean 4.06, SD 1.18).

XSL•FO
RenderX

Table 1. Start2Cycle app page views.

Page	Views (absolute), n	Views (relative), %
Coach	1159	26.9
Competition	578	13.4
Route	547	12.7
Team	539	12.5
Routes	529	12.3
Track	507	11.8
Personal	268	6.2
Trophies	119	2.8
Bug report	58	1.3
Total	4304	100

After data cleaning, the log data showed that the participants averaged 37 app use sessions (SD 29.35), with an average of 5 actions (SD 5.86) per session during the field trial. The high SDs indicate substantial variability in activity logging between the participants.

Table 1 depicts the number of views per app page during the field trial. It is clear that the *Coach* page was visited the most, although it has to be noted that this was the default starting page for a session. The table further demonstrates that the gamification features on both the *Competition* and *Team* page were frequently viewed, as well as the route tracking features (*Track*, *Route*, and *Routes* pages). This is straightforward, as without users tracking routes, no gamification could be enabled. The *Personal* and *Trophies* page were not often frequented by the users. The same applies for the *Bug report* page.

Markov Chain Analysis

Table 2 demonstrates the transition matrix of the Markov Chain, and Figure 2 visualizes the transitions between the app pages. Not all transitions are displayed in Figure 2. A cut-off of a 0.13 probability was used to improve readability. For the *Trophies* and *Personal* page, the highest incoming probability is also displayed to improve readability and logic of the visualization. Without these incoming probabilities, the pages would be visually isolated in the figure as all incoming edges have a probability lower than 0.13.

Figure 2 demonstrates that the majority of the sessions (76.9%, 634/824) start from the *Coach* page, which was intended by design. The *Coach* page, thus, served as the point from which most of the users started actions from in the app (yellow lines in Figure 2) and returned to before exiting the app. The remaining 23.1% (190/824) of sessions start from other pages than the *Coach* page. This is because of users not exiting the app (eg, switching to another app) and then reopening the app on the same page they left off, most often the *Track* and *Competition* page.

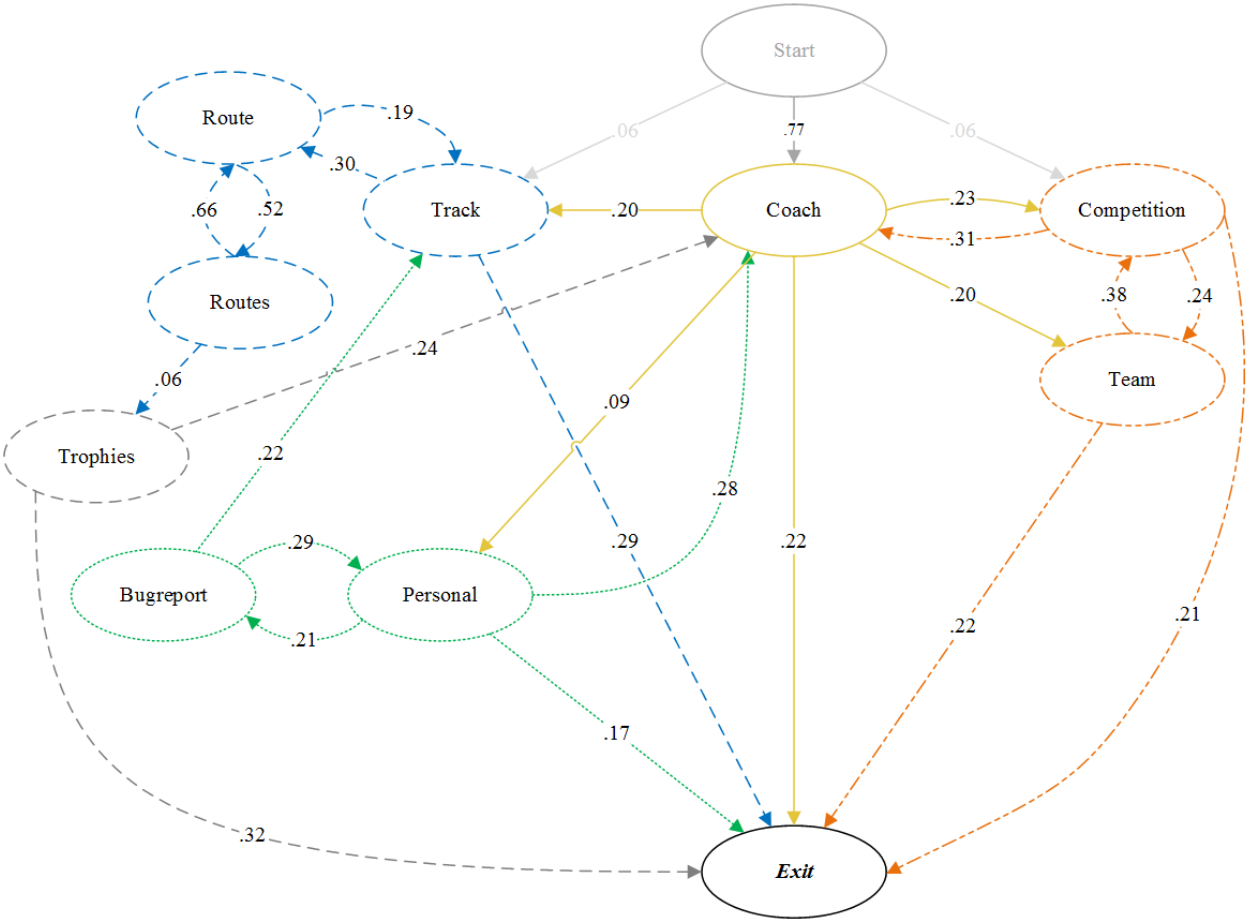
Table 2. Transition matrix (probability of transitioning from page A [row] to B [column]). Rows sum up to 1.

	Trophies	Bug report	Coach	Competition	Personal	Route	Routes	Team	Track	Exit
Trophies	__[a]	—	0.24	0.11	0.04	—	0.08	0.12	0.08	0.32
Bug report	—	—	0.12	0.09	0.29	—	0.05	0.12	0.22	0.10
Coach	0.03	—	—	0.23	0.09	0.00	0.04	0.20	0.20	0.22
Competition	0.05	—	0.31	—	0.05	0.01	0.09	0.24	0.05	0.21
Personal	—	0.21	0.28	0.04	—	0.00	0.08	0.13	0.09	0.17
Route	—	—	0.13	—	0.04	—	0.52	0.01	0.19	0.11
Routes	0.06	—	0.07	0.03	0.03	0.66	—	0.05	0.04	0.07
Team	0.04	—	0.11	0.38	0.08	0.00	0.11	—	0.06	0.22
Track	0.01	—	0.13	0.03	0.05	0.30	0.07	0.12	—	0.29
Exit	—	—	—	—	—	—	—	—	—	—

[a]No transitions occurred between the two corresponding app pages.

XSL•FO
RenderX

Figure 2. Start2Cycle navigation visualization.



Visual inspection of the transitions between the different app pages roughly exposes 3 *Start2Cycle* app session types. A first session type can be labeled a *gamification* session (orange striped lines). After launching the app, users would then go from the *Coach* page to either the *Competition* page, to see whether or not their team is still ahead of the other in terms of cycled kilometers (23%) or check how they are ranked within their own team (on the *Team* page, 20%). Switching between these 2 pages is a likely next action. These sessions are often followed by a return to the *Coach* page, particularly from the *Competition* page (31%), or the app session is ended.

Second, a *route tracking* session can be distinguished (blue striped lines in Figure 2). In such a session, a user would start the app with the intention of logging a cycling activity. There is a 20% probability that a user would start the app with this action in mind from the *Coach* page. After tracking a route, there is a high probability that the user will check the route they rode (on the *Route* page) or other routes they have previously cycled (on the *Routes* page) and switch back and forth between these 2 pages. Eventually, the user is likely return to the *Coach* page (13%) and leave his session or perform new actions from there. The app is often left from the *Track* page (29%). This is likely because of the fact that an inactivity of longer than 30 min occurred because of cycling or because of errors in route tracking in the early phases of the trial (a known problem of the app), which prompted the user to leave the app. As could be derived from the limited number of page views displayed in Table 1, the *Trophies* page (grey striped lines), where a user could check whether he or she unlocked a new video in the app, has very limited incoming transitions, the highest being an incoming transition from the *Routes* page. After checking the *Trophies* page, there is a high probability the user will either leave the app (32%) or proceed to the *Coach* page (24%).

The last session type is labeled a *bug report* session (green dotted lines). These sessions were not as frequent as there is only a 9% probability that a user would go from the *Coach* page to his personal page and subsequently report an error or malfunction in the app. When reporting these tracking errors, users often immediately switched to the *Track* page, most likely for a new attempt to track their activity.

Sequence Clustering

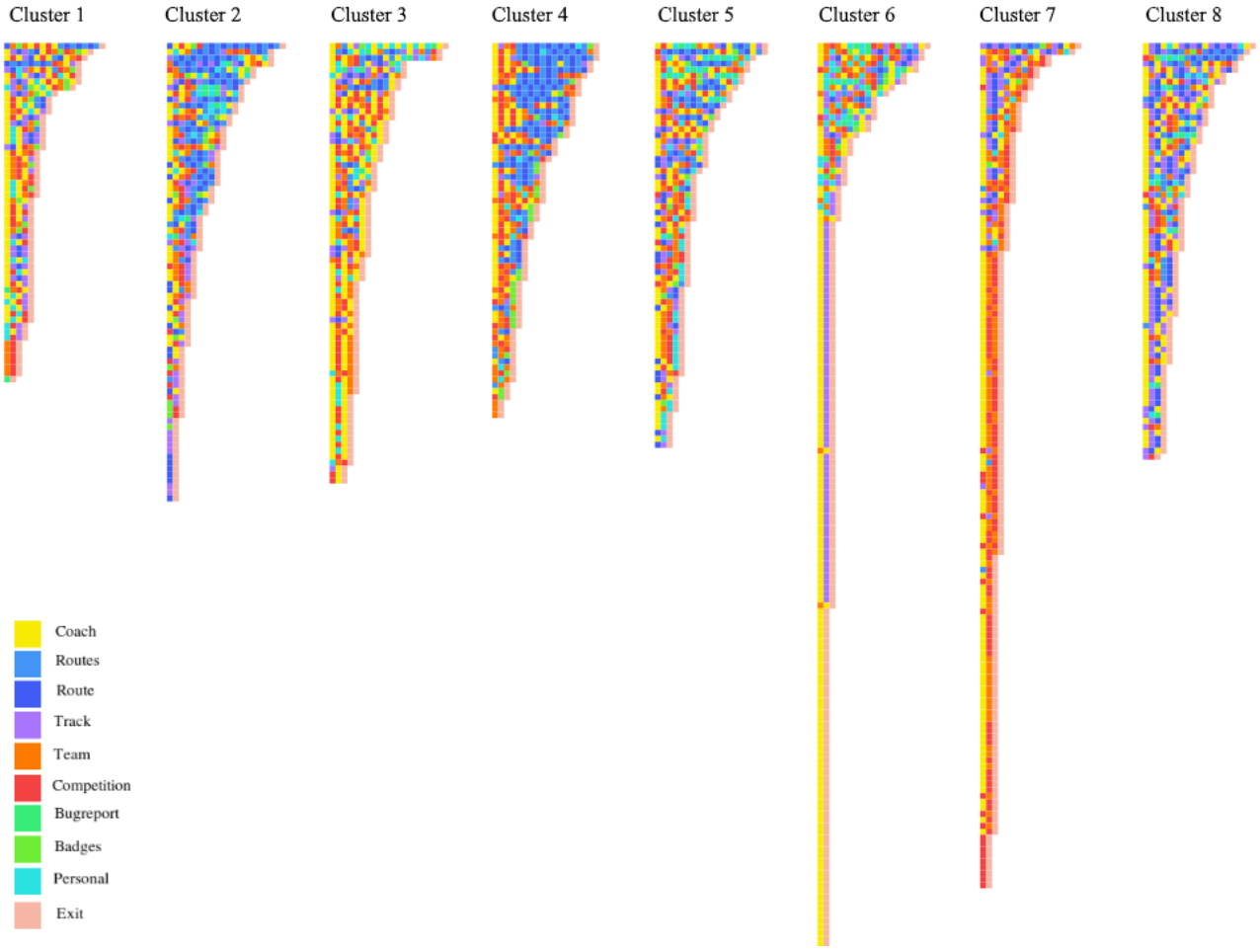
Although 3 types of user sessions were identified through visual inspection, the actual sequences in the data are more complex. As an example, a session could be considered a *route tracking* session when use of the *Track* page and consequent use of, for example, the *Route* and *Routes* page occurs, but in reality, a

XSL•FO
RenderX

user is likely to use gamification features or the coach multiple times in the same sessions. Nevertheless, applying the sequence clustering algorithm with varying *K* clusters, distinct clusters containing specific session types, could be observed. The results of the algorithm are an assignment of each sequence (or user's session) to 1 of *K* clusters and first-order Markov Chains, 1 for each cluster. Figure 3 displays the assignment of the sessions in *K*=8 clusters.

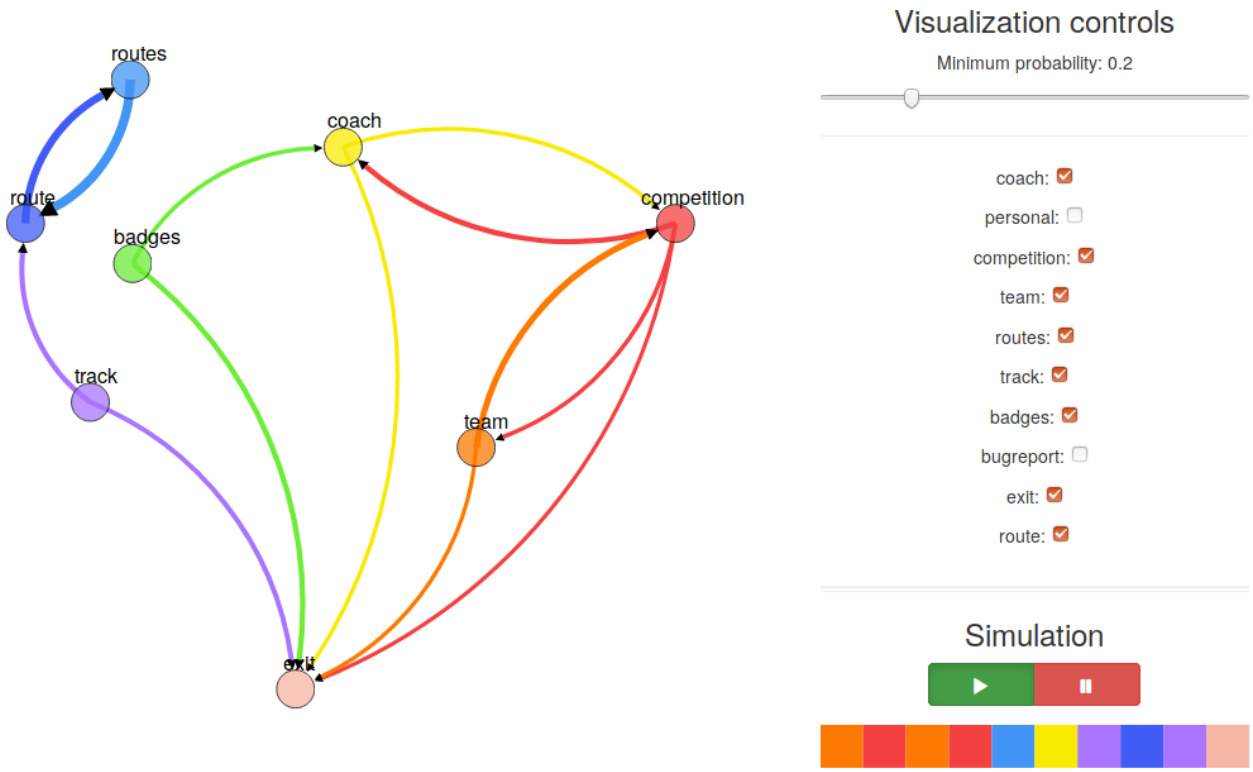
Only sequences with a length between 2 and 20 actions are depicted in the figure to avoid cluttering. Each row in the visualization in Figure 3 corresponds to a session, and each cell in the row corresponds to a certain action made by the user. Although there is an overlap between clusters, particularly for longer sessions with a varying number of different actions, the tails of the clusters consist of shorter sessions with clear purposes. Cluster 7 shows clear abstraction of gamification sessions. These sessions are typically short and consist of a user opening the app, checking the leaderboards, and exiting the app. Furthermore, cluster 6 makes a distinct grouping of presumed erroneous route tracking sessions consisting of the sequence *Coach-Track-Exit*. Clusters 2 and 4, on the other hand, distinctively contain more sequences involving checking of the *Route* and *Routes* pages.

Figure 3. Visualization of sequence clustering.



XSL•FO
RenderX

Figure 4. Screenshot of the visualization tool.



Web-Based Visualization Tool

Figure 4 presents a screenshot of the visualization and simulation tool that was developed to allow researchers to dynamically explore the *Start2Cycle* data. In the tool, researchers can drag around selected nodes, explore transitions from given probability cut-offs, and run simulated paths through the *Start2Cycle* app.

Discussion

Principal Findings

The purpose of this study was to demonstrate how data mining can be applied to evaluate the development of mHealth apps. It is clear that by performing a Markov Chain analysis on the log data of the *Start2Cycle* app, it was possible to identify different types of user sessions, to assess the popular features of the app, and cluster frequently occurring app usage patterns.

In the light of further development of the app, this type of analysis has provided insights that may not have been discerned using traditional methods such as usability tests, evaluation surveys, or expert reviews, particularly when dealing with apps with a high number of app pages. Briefly stated, using Markov Chains affords mHealth developers and researchers to assess whether the use of, and the navigation in their app, happens as they envisioned it. In our analysis, 2 distinct types of app usage sessions could be identified, which aligned with the envisioned use of the app: a *gamification* session and a *route tracking* session. Evidently, the *Start2Cycle* app is not excessively complex, so use of the app could be expected to be straightforward. However, this was not entirely the case.

It was clear that the gamification features involving competition and cooperation within and between teams included in the app were appealing to the users, which was expected based on the review of literature on gamification [34,35]. Users often checked these gamification pages and navigated smoothly from one to another, which could be derived from the transition probabilities calculated in the Markov Chain analysis. On the other hand, the trophies concept was clearly less appreciated despite being easily accessible through the navigation menu. The Markov Chain analysis showed little incoming navigation to the *Trophies* page. It appears that the concept behind this feature should be improved to make it more appealing. Anecdotal conversations with some of the respondents revealed that the video content did not appeal to the users, and more tangible trophies (eg, coupons, water bottles, or cycling jerseys) or virtual badges were preferred. Previous research on the effectiveness of badges has demonstrated mixed results [36]. Further research is required to assess the effectiveness in the particular context of this study.

Furthermore, Markov Chain–based sequence clustering demonstrated how similar sequences of actions can be grouped together in clusters. This method becomes particularly interesting when apps contain significantly more pages than the *Start2Cycle* app used in this study, and interpretation from matrices and visualizations is less straightforward. Furthermore, by using this technique, a participant's app use profile can automatically be determined based on the types of sessions performed throughout the use of the app. The user interface of the app, for example, could then be tailored specifically to this profile.

The application of Markov Chains is only 1 method within a broader range of methods in the context of data mining. The

XSL•FO
RenderX

vast amounts of data generated by mHealth apps and games are still largely unexplored. This study aimed to illustrate the potential of 1 method to generate useful insights from log data that can foster further and improved development of mHealth apps and games. Data mining can advance behavior change research by uncovering usage patterns and insights that could not be uncovered through traditional methods.

Limitations

Some limitations of this study should be noted. The aim was to demonstrate how data mining, Markov Chain analysis in particular, can be applied in the development of mHealth apps. The data used in this study originate from a field trial with primarily usability testing purposes. Therefore, certain properties of the sample may influence how the app was used. For example, although the app is designed to motivate adults to start cycling or to cycle more often, some of the participants in this study were already highly involved in active cycling. It was opted to work with active cyclists to ensure that sufficient (activity) log data would be generated during the field trial. This implies that the sample used in this study is presumably not representative of a population that would typically be asked to use the app. However, as the purpose of this study is primarily to demonstrate the potential of Markov Chain analysis in the development of mHealth apps, the characteristics of the sample can be considered somewhat less important.

Second, it was chosen to use first-order Markov Chains, which assume that the probability of moving to a next state depends only on the current state (and thus not on previous states). This benefits the intuitiveness and simplicity of our analysis. A first-order Markov Chain does not change, although the representation of a higher-order Markov Chain would have to be conditioned on the previous transitions. Clearly, this would make the proposed visualization much less comprehensible. Furthermore, a substantial amount of data is needed to construct a Markov Chain of order $m>1$ that is a good representation of the global user behavior. If the chosen m is too high, given the number of samples, then most history state sequences of length m do not occur in the data, or only occur infrequently, which causes the transition matrix, conditioned on that history sequence, to be constructed from insufficient data.

Finally, the user sessions used in the analyses were obtained using a heuristic based on time between actions in the app. Better would have been to collect a session ID along with every user session. This was overlooked in the design of the app and the study. Therefore, cut-offs between session may not be entirely accurate. It is recommended for future studies that session IDs are collected along with the data.

Future Research

Further research should build on the insights of this study and explore how data mining can help to improve mHealth behavior change research. Other data mining techniques are available, and studies using machine learning algorithms, that is, predictive modeling, are strongly encouraged as these could help predicting user actions in the app using different input variables such as user profile variables or placement of different elements on a page. If an accurate model of app usage can be constructed, then the parameters (such as the placement of an element) can be manipulated to maximize the probability that a user takes a certain action. Furthermore, identified usage patterns could also help predicting user attrition and consequently allow an app to take appropriate actions. For example, detecting that a user is likely to drop out of an intervention by his usage pattern can trigger an app to engage the user again at an appropriate moment. However, such analyses would require a larger number of participants to be monitored for an extensive amount of time to detect whether or not they drop out. Moreover, user attrition does not necessarily imply that the participant failed to adopt a healthier lifestyle. It may well be that he or she has no need for an app to support behavior change any longer. In this study, data from 22 participants was used and was sufficient for obtaining insightful results. We recommend higher sample sizes, however.

It should be noted that evidently there is a limit to what can be learned through automated analyses. Additional use of traditional quantitative and qualitative techniques such as surveys and interviews will yield a more complete picture of how the user evaluates the app. In this study, routes were often not tracked at all, the tracking screen froze, all data were lost, or the tracking was fine but the kilometers were not added to the total team score, which could be frustrating to competitive users. Reports of such bugs are best detected using evaluation surveys. Furthermore, correct data selection and cleaning can be a time-consuming and difficult undertaking, and a health researcher may not always be trained to do this.

Conclusions

This study demonstrated the potential of data mining in development and evaluation of mHealth apps. Results demonstrate how health researchers and developers in the field of mHealth can use and learn from new data analysis methods originating from the field of data mining. Our study shows how usage patterns of mHealth apps can be revealed and provide new insights to how interventions are used. As such, these insights can lead to further improvement of mHealth interventions and consequently improve their effectiveness.

A plea for interdisciplinary research and collaboration with data scientists is therefore required. In this regard, development of visualization and simulation tools that make data mining methods and results more accessible to researchers and developers with limited data mining skills is highly encouraged.

Availability of Data and Materials

All experiments and calculations in this study were performed using Python 3. All data and code to reproduce the experiments performed in this study, or to create the proposed visualization based on your app logs, can be found on our GitHub page [37,38] under an open license for noncommercial use.

XSL•FO
RenderX

Acknowledgments

This study originates from the CONAMO project, funded by the Imec ICON research program, with project support of Vlaio. Companies and organizations involved in the project are IMEC, imec-IDLab, imec-MICT, imec-SMIT, imec-Living Labs, UGent-Victoris, Vlaamse Radio- en Televisieomroeporganisatie (VRT), Energy Lab, and Rombit. GV is funded by a PhD SB fellow scholarship of FWO (1S31417N).

The authors express their special gratitude to public broadcaster and project partner VRT and its team for the development of the app, especially Matthias De Vriendt for outlining the raw data structure and format.

Conflicts of Interest

None declared.

References

1. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. J Am Med Assoc 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
2. Poushter J, Stewart R. Pew Research Center. 2016. Smartphone ownership and internet usage continues to climb in emerging economies URL: https://www.pewresearch.org/wp-content/uploads/sites/2/2016/02/pew_research_center_global_technology_report_final_february_22__2016.pdf [accessed 2019-05-06] [WebCite Cache ID 78ABvSvPi]
3. Aitken M, Clancy B, Nass D. IQVIA. 2017. The Growing Value of Digital Health: Evidence and Impact on Human Health and the Healthcare System URL: https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health [accessed 2019-05-06] [WebCite Cache ID 78ACGVFcG]
4. Bateman DR, Srinivas B, Emmett TW, Schleyer TK, Holden RJ, Hendrie HC, et al. Categorizing health outcomes and efficacy of mHealth apps for persons with cognitive impairment: a systematic review. J Med Internet Res 2017 Dec 30;19(8):e301. [doi: 10.2196/jmir.7814] [Medline: 28855146]
5. Conway N, Campbell I, Forbes P, Cunningham S, Wake D. mHealth applications for diabetes: user preference and implications for app development. Health Informatics J 2016 Dec;22(4):1111-1120. [doi: 10.1177/1460458215616265] [Medline: 26635324]
6. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
7. Müller AM, Alley S, Schoeppe S, Vandelanotte C. The effectiveness of e-& mHealth interventions to promote physical activity and healthy diets in developing countries: a systematic review. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109 [FREE Full text] [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
8. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. Adv Nutr 2017 May;8(3):449-462. [doi: 10.3945/an.116.014100] [Medline: 28507010]
9. Clawson J, Pater J, Miller A, Mynatt E, Mamykina L. No longer wearing: investigating the abandonment of personal health-tracking technologies on craigslist. 2015 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 07-11, 2015; Osaka, Japan p. 647-658. [doi: 10.1145/2750858.2807554]
10. Lazar A, Koehler C, Tanenbaum J, Nguyen D. Why we use and abandon smart devices. 2015 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 07-11, 2015; Osaka, Japan p. 635-646.
11. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
12. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
13. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 07;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
14. Spanakis G, Weiss G, Boh B, Lemmens L, Roefs A. Machine learning techniques in eating behavior e-coaching. Pers Ubiquit Comput 2017 Jun 8;21(4):645-659. [doi: 10.1007/s00779-017-1022-4]
15. Virtanen S, Rost M, Higgs M, Morrison A, Chalmers M, Girolami M. Non-parametric Bayes to infer playing strategies adopted in a population of mobile gamers. Stat 2015;4(1):46-58.
16. Andrei O, Calder M, Higgs M, Girolami M. Probabilistic Model Checking of DTMC Models of User Activity Patterns. In: Norman G, Sanders W, editors. Quantitative Evaluation of Systems. Cham: Springer; 2014.

XSL•FO
RenderX

17. Hu Z, Ma Y, Mei Q, Tang J. Arxiv.org. 2017. Roaming across the Castle Tunnels: an Empirical Study of Inter-App Navigation Behaviors of Android Users URL: https://arxiv.org/abs/1706.08274 [accessed 2019-05-06] [WebCite Cache ID 78ADgec5A]
18. Harpstead E, Zimmermann T, Nagapan N, Guajardo J, Cooper R, Solberg T. What drives people: Creating engagement profiles of players from game log data. 2015 Presented at: Annual Symposium on Computer-Human Interaction in Play; October 05-07, 2015; London, United Kingdom p. 369-379.
19. Andrei O, Calder M, Chalmers M, Morrison A, Rost M. Probabilistic Formal Analysis of App Usage to Inform Redesign. 2016 Presented at: 12th International Conference on Integrated Formal Methods; June 01-05, 2016; Reykjavik, Iceland. [doi: 10.1007/978-3-319-33693-0_8]
20. Arden-Close EJ, Smith E, Bradbury K, Morrison L, Dennison L, Michaelides D, et al. A visualization tool to analyse usage of web-based interventions: the example of positive online weight reduction (POWeR). JMIR Hum Factors 2015 May 19;2(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.4310] [Medline: 27026372]
21. Miller S, Ainsworth B, Yardley L, Milton A, Weal M, Smith P. A Framework for Analyzing and Measuring Usage and Engagement Data (AMUsED) in Digital Interventions: Viewpoint. Journal of Medical Internet Research 2019;21(1):e10966. [doi: 10.2196/10966]
22. Borges J, Levene M. Evaluating variable-length Markov chain models for analysis of user web navigation sessions. IEEE Trans Knowl Data Eng 2007 Apr;19(4):441-452. [doi: 10.1109/TKDE.2007.1012]
23. Becker M, Singer P, Lemmerich F, Hotho A, Helic D, Strohmaier M. Photowalking the City: Comparing Hypotheses About Urban Photo Trails on Flickr. In: Liu TY, Scollon C, Zhu W, editors. Social Informatics. Cham: Springer; 2015.
24. Singer P, Helic D, Taraghi B, Strohmaier M. Detecting memory and structure in human navigation patterns using Markov chain models of varying order. PLoS ONE 2014 Jul 11;9(7):e102070. [doi: 10.1371/journal.pone.0102070]
25. 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]
26. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
27. Staiano AE, Abraham AA, Calvert SL. Motivating effects of cooperative exergame play for overweight and obese adolescents. J Diabetes Sci Technol 2012 Jul 01;6(4):812-819 [FREE Full text] [doi: 10.1177/193229681200600412] [Medline: 22920807]
28. Marker AM, Staiano AE. Better together: outcomes of cooperation versus competition in social exergaming. Games Health J 2015 Feb;4(1):25-30 [FREE Full text] [doi: 10.1089/g4h.2014.0066] [Medline: 26181677]
29. Page L, Brin S, Motwani R, Winograd T. Stanford InfoLab. 1999. The PageRank citation ranking: Bringing order to the web URL: http://ilpubs.stanford.edu/422/ [accessed 2019-05-06] [WebCite Cache ID 78AEN9bAJ]
30. Ferreira D, Zacarias M, Malheiros M, Ferreira P, editors. Approaching Process Mining with Sequence Clustering: Experiments and Findings. In: Alonso G, Dadam P, Rosemann M, editors. Business Process Management. Berlin, Heidelberg: Springer; 2007.
31. Cadez I, Heckerman D, Meek C, Smyth P, White S. Model-based clustering and visualization of navigation patterns on a web site. Data Min Knowl Discov 2003;7(4):399-424. [doi: 10.1023/A:102499261]
32. Moon TK. The expectation-maximization algorithm. IEEE Signal Process Mag 1996;13(6):47-60. [doi: 10.1109/79.543975]
33. Bostock M, Ogievetsky V, Heer J. D³ data-driven documents. IEEE Trans Vis Comput Graph 2011;17(12):2301-2309. [doi: 10.1109/TVCG.2011.185]
34. Alahäivälä T, Oinas-Kukkonen H. Understanding persuasion contexts in health gamification: a systematic analysis of gamified health behavior change support systems literature. Int J Med Inform 2016 Jan;96:62-70. [doi: 10.1016/j.ijmedinf.2016.02.006] [Medline: 26944611]
35. Sailer M, Hense J, Mayr S, Mandl H. How gamification motivates: an experimental study of the effects of specific game design elements on psychological need satisfaction. Comput Human Behav 2017 Apr;69:371-380 [FREE Full text] [doi: 10.1016/j.chb.2016.12.033]
36. Hamari J. Do badges increase user activity? A field experiment on the effects of gamification. Comput Human Behav 2017 Jun;71(6):469-478. [doi: 10.1016/j.chb.2015.03.036]
37. GitHub. URL: https://github.com/IBCNServices/MCAppAnalysis [accessed 2019-05-10] [WebCite Cache ID 78HPnWvXw]
38. GitHub. URL: https://ibcnservices.github.io/MCAppAnalysis/ [accessed 2019-05-10] [WebCite Cache ID 78HPwqdY9]

Abbreviations

DBCIs: digital behavior change interventions
ID: identifier
mHealth: mobile health

XSL•FO
RenderX

Edited by M Focsa; submitted 16.08.18; peer-reviewed by M Pobiruchin, O Perski; comments to author 30.10.18; revised version received 24.01.19; accepted 14.03.19; published 07.06.19.

Please cite as:
Stragier J, Vandewiele G, Coppens P, Ongenae F, Van den Broeck W, De Turck F, De Marez L
Data Mining in the Development of Mobile Health Apps: Assessing In-App Navigation Through Markov Chain Analysis
J Med Internet Res 2019;21(6):e11934
URL: https://www.jmir.org/2019/6/e11934/
doi:10.2196/11934
PMID:31237838

©Jeroen Stragier, Gilles Vandewiele, Paulien Coppens, Femke Ongenae, Wendy Van den Broeck, Filip De Turck, Lieven De Marez. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Use of Web-Based Health Services in Individuals With and Without Symptoms of Hypochondria: Survey Study

Christiane Eichenberg[1], Prof Dr; Markus Schott[1], MSc

Sigmund Freud Privat Universität Wien, Vienna, Austria

Corresponding Author:
Markus Schott, MSc
Sigmund Freud Privat Universität Wien
Freudpl 3
Vienna, 1020
Austria
Phone: 49 15124049991
Email: markus.s.c.schott@gmail.com

Abstract

Background: An increasing number of people consult physicians because of distressing information found online. Cyberchondria refers to the phenomenon of health anxiety because of online health information.

Objective: This study aimed to examine online health research of individuals with and without symptoms of hypochondria and their impact on health anxiety as well as behavior.

Methods: An online survey was conducted. Demographic data, health-related internet use, and general health behavior were assessed. The illness attitude scale was used to record symptoms of hypochondria.

Results: The final sample consisted of N=471 participants. More than 40% (188/471) of participants showed at least some symptoms of hypochondria. Participants with symptoms of hypochondria used the internet more frequently for health-related purposes and also frequented more online services than individuals without symptoms. Most online health services were rated as more reliable by individuals with symptoms of hypochondria. Changes to behavior such as doctor hopping or ordering nonprescribed medicine online were considered more likely by individuals with symptoms of hypochondria.

Conclusions: Results show that individuals with symptoms of hypochondria do not turn to online research as a result of lacking alternatives but rather consult health services on- as well as offline.

(J Med Internet Res 2019;21(6):e10980) doi:10.2196/10980

KEYWORDS

hypochondria; eHealth; anxiety; survey

Introduction

Background

Currently, the internet provides an enormous amount of information about a range of health topics. According to a representative study with N=2411 participants [1], 63.5% of German internet users rely on the internet as a guide concerning health-related issues. In addition to traditional media such as television, radio, and print, the internet has become an important source of information related to health issues. Information found online does not affect personal health behavior as much as a doctor's or family member's recommendation; however, their importance is comparable with the influence of a pharmacist [1]. According to a US representative telephone survey [2], 86% of online health users reported that online research had successfully met their information needs and also regarded the information found as reliable.

Online health research may lead to positive and preventative activities such as exercise, healthier eating habits, improved adherence to medication, and empowered health decisions [3]. However, these possibilities present a problem when used as diagnostic tools by laypersons [4]. As the internet provides access to a large body of information, it is a crude means of self-diagnosis. For example, multiple medical possibilities are typically presented in response to internet searches of medical concerns [5]. Therefore, searching for medical information on the internet has the potential to lead to even greater levels of uncertainty and consequently exacerbate health anxiety for vulnerable individuals. This might aggravate already existing

XSL•FO
RenderX

worries about being sick with a certain disease for patients suffering from hypochondria.

Definition

Hypochondria is a debilitating disorder that occurs in about 1 to 5% of people once in a lifetime [6]. The Diagnostic and Statistical Manual of Mental Disorders (DSM), 4th edition (Text Revision) defines this disorder as a somatoform disorder [7]. The newly published DSM 5th edition replaces the diagnosis of hypochondria with the diagnoses of somatic symptom disorder and illness anxiety disorder [8]. Hypochondria refers to a persistent preoccupation with, or fears about the possibility of having or developing 1 or more serious progressive or life-threatening diseases. The preoccupation is associated with a hypervigilance to and catastrophic misinterpretation of bodily signs or symptoms including normal sensations and is accompanied by avoidance or repetitive behaviors. Minor symptoms are thus interpreted as evidence of suspected diseases, and concerns persist despite reassurance by physicians [7,9]. Disease conviction can be maintained by repetitive symptom-checking or reassurance seeking including activities on the internet: retrieving medical information or entering self-observed symptoms into online diagnostic systems. In addition, sharing information online with other affected people may be an attempt to get rid of unpleasant thoughts and feelings concerning their own state of health.

This more recent phenomenon has been referred to as cyberchondria [10]. Cyberchondria has been defined as an excessive search on the internet for health-related information, which is driven by an underlying need to ease anxiety, but subsequently results in symptoms worsening [11]. It is a form of reassurance-seeking behavior. Instead of obtaining support via online interactions with similarly worried individuals, those with cyberchondria experience their health anxiety as exacerbated, mainly because of new pathologies found online that might trigger new worries [2]. Although current research already declared the phenomenon of cyberchondria as a distinct clinical disorder [10], the prevailing view is that cyberchondria is part of hypochondria/health anxiety. Unfortunately, little research has been conducted in this area. There have only been few elaborate studies on internet use and health fears to date.

Available Research

A survey of over 500 respondents by White and Horvitz [10] found that more than 60% of participants repeatedly searched for a specific health concern in multiple online sessions. In a further analysis of thousands of online interaction logs, it was found that over an 11-month period, 13.5% of study participants entered the exact same health-related terms into a search engine on more than 1 occasion [10]. Starcevic and Berle [11] suggested that these excessive searches only serve to reinforce a person's original anxiety. Over a third of the sample (n=198, 38.4%) said they experienced an increase in anxiety following an online search for health information. The majority (n=365, 70.7%) of participants in this study believed that the fear was caused by the content of websites visited as part of their research efforts; that is, navigation escalation. A total of 7 out of 10 respondents were still researching serious illnesses for weeks and months after a conducted search, suggesting that these escalated search sessions had a lasting impact.

Aside from causing unwarranted levels of worry and distress, there may also be economic costs to cyberchondria [12]. Even though no studies have examined the costs directly associated to online health searches to this date, there is evidence that those who are especially health anxious constitute a considerable economic burden to society. For example, health care costs and losses in productivity attributable to medically unexplained symptoms account for an estimated £3 billion in 2008 in the United Kingdom alone [13]. Cyberchondria may be likely to be responsible for a significant proportion of this amount. In an online study (N=240), Eastin and Guinsler [14] found a moderating effect of health anxiety on the link between health-related online research and the use of health care. Results highlighted that with increasing health anxiety, the association between the extent of online information search and appointments with physicians became stronger. Thus, health anxiety seems to be an important determinant of how online medical research affects the decision to visit a doctor. Furthermore, an analysis of anonymous search logs showed that those who frequently searched for health information online ended their search sessions with queries about local health care services [15]. In contrast, research has also suggested that cyberchondria can lead to deterioration within the doctor-patient relationship [16], which may lead to further health care costs (eg, visits to multiple doctors). Therefore, it is important to better understand this disorder to inform new strategies to minimize its negative consequences.

Objectives

This study aimed to examine usage patterns of online health research and their impact on health anxiety and behavior. It was hypothesized that individuals with symptoms of hypochondria would use different online health services more often than individuals without symptoms of hypochondria. Furthermore, it was hypothesized that individuals with symptoms of hypochondria would evaluate these services with regard to information quality better than individuals without symptoms of hypochondria. We also investigated how participants react to health-related online research and how they rated their potential to alleviate anxiety.

Methods

Study Design

The study was carried out as a Web-based survey. To generate a sample as heterogeneous as possible, the questionnaire was distributed to more than 180 German-speaking health forums, ranging from very specific health-related forums (eg, online self-help for alcoholics) to more general ones. Information about the purpose and the course of the study alongside an invitation link to the online questionnaire were published in the forums. The data collection period lasted 4 weeks. Completing the questionnaire took about 10 to 15 min.

Ethical approval was obtained from the Sigmund Freud University Vienna ethics committee (QBABQXGPA@APWF87082).

XSL•FO
RenderX

Statistical analyses were conducted by using ASW Statistics for Windows, Version 18.0 by SPSS Inc.

Material

The online questionnaire consisted of 2 parts: Part I was self-constructed and included demographic data (5 items), health-related internet use (11 items), and general health behavior (7 items). Part II was standardized and assessed existing symptoms of hypochondria.

A pretest was carried out in a specially selected health forum. The N=17 returns were analyzed, and the instrument was revised regarding its practicability, comprehensibility, and completeness of item formulation.

Demographic Data, Internet Use, and Health Behavior

In addition to age, gender, and nationality of participants, the highest level of education and subjects' current occupational status were assessed.

To acquire information about the usage of health-related online services, participants were asked to report their level of internet literacy, the frequency of internet use, as well as internet services used for health-related research. It was also recorded how subjects perceived the quality of various online health services and whether participants experienced effects of different online health services on one's own health anxiety.

Study participants indicated the availability of adequate medical care in the immediate vicinity on a 5-point Likert scale ("very good" to "deficient"), the frequency of doctor visits, and number of different doctors visited within the last year.

Illness Attitude Scale

The Illness Attitude Scale (IAS) was used to record health anxiety. This instrument consists of 2 scales (1st disease anxiety, 2nd disease behavior) with 29 items on various aspects of hypochondria to be answered on a 5-step scale. A total of 9 scales are formed from 3 questions each. In addition, 2 additional questions are included. The IAS has 9 subscales: (1) worry about illness, (2) concerns about pain, (3), health habits, (4) hypochondriacal beliefs, (5) thanatophobia (fear of death), (6) disease phobia, (7) bodily preoccupations, (8) treatment experience, and (9) effects of symptoms. Each item is rated on a 0 to 4 Likert scale, and 27 of the 29 items are used in the total score, which ranges from 0 to 108. Findings show that a cut-off of 47 points yields a sensitivity of 96% and specificity of 95%. The Illness Attitude Scale has a retest-reliability between 0.89 and 0.93 and a Cronbach's Alpha of $\alpha = 0.90$ [17].

Results

Sample

The final sample consisted of N=471 participants recruited from various health forums on the internet, of which 397 (84.3%) were female. Participants were aged between 14 and 84 years (mean 40.0, SD 13.25). Most participants were in a relationship or married (310/471, 65.8%). At the time of the survey, more than half of participants (273/471, 57.9%) had finished school with at least a high school diploma.

Illness Attitude Scale

Results indicated that more than a quarter of participants showed at least some symptoms of hypochondria (70/471, 14.9%) or could even be diagnosed with hypochondria (120/471, 25.5%). To ease statistical analyses and to create more balanced group sizes, these 2 categories were merged into 1 (190/471, 40.4%).

Comparison Between Participants With and Without Hypochondria

Internet Use

Participants used a wide variety of services on the internet on a fairly regular basis (Table 1). Consequently, it is not surprising that a majority of subjects rated their own internet literacy as "advanced" (334/471, 70.9%) or even "expert" (67/471, 14.3%).

Half of the sample reported that they had used the internet more than 10 times in the past year to search for acute (215/471, 45.7%) or chronic symptoms (217/471, 46.2%). As expected, results showed that individuals with symptoms of hypochondria had been searching for their own acute (χ^2_4=28.82; n=400; P<.001; phi=0.27) and chronic symptoms χ^2_4=28.167; n=400; P<.001; phi=0.27) significantly more often than those with low health anxiety. For example, 76% (144/190) of individuals with symptoms of hypochondria searched for symptoms more than 10 times on the internet last year compared with only 36.3% (n=101/281) of individuals without symptoms of hypochondria.

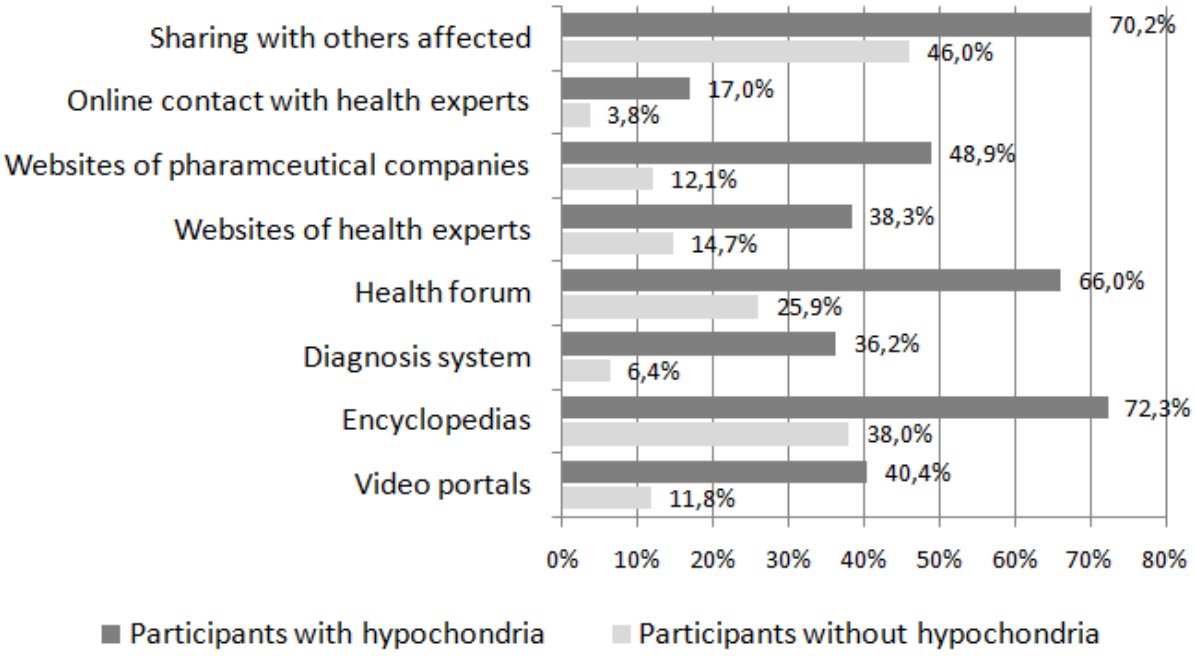
Individuals with symptoms of hypochondria also differed significantly in the types of online health services used from individuals without symptoms of hypochondria (all differences P<.001; see Figure 1). In particular, health services to share with others affected and health forums were used several times a day by 35% (66/281) and 50% (95/281) of individuals with symptoms of hypochondria, respectively, whereas only 12.6% (35/281) of individuals without symptoms of hypochondria used these services on a daily basis.

Table 1. Intensity of use of various internet services (N=470).

Reason to use	Frequency			
	At least once a day, n (%)	At least once a week, n (%)	At least once a month, n (%)	Never, n (%)
Professional	198 (42.2)	92 (19.6)	63 (13.4)	115 (24.5)
Entertainment	282 (60.1)	119 (25.3)	23 (4.9)	22 (4.7)
Social media	295 (62.8)	130 (27.7)	17 (3.6)	28 (6)
Information	323 (68.7)	133 (28.5)	10 (2.2)	3 (0.6)

XSL•FO
RenderX

Figure 1. Weekly or more often usage of different online services for individuals with and without symptoms of hypochondria.



Similarly, individuals with symptoms of hypochondria used a greater variety of services than individuals without symptoms of hypochondria (t_{358}=–4.06; P<.001; d=0.65). Individuals with symptoms of hypochondria consulted on average mean 5.72 (SD 1.79) various health services on the internet, whereas individuals without these symptoms only visited mean 4.51 (SD 1.93) services.

Evaluation of Online Information Quality

In general, study participants' appraisal of online health information was critical. The health service with the highest rating was only considered to be "rather reliable." The quality of online help by experts (mean 1.97, SD 0.81) was rated highest. However, exchange with medical laypersons (mean 1.77, SD 0.81) was considered to be almost as reliable. In contrast, information from pharmaceutical companies (mean 1.26, SD 0.86) was assessed more critically, similar to online diagnostic systems (mean 0.82, SD 0.79) and information on video portals (mean 0.53, SD 0.71).

Unsurprisingly, most online health services were rated as significantly more reliable by individuals with symptoms of hypochondria than by individuals without symptoms of hypochondria (see Table 2).

XSL•FO
RenderX

Table 2. Evaluation online information quality.

Online services	Mean (SD)	*t* (*df*)	Cohen *d*	*P* value
Online contact with experts	__[a]	–0.53 (213)	0.10	.59
Symptoms of hypochondria	2.04 (0.88)	—	—	—
No symptoms of hypochondria	1.95 (0.80)	—	—	—
Sharing with others affected	—	–2.26 (363)	0.36	.02
Symptoms of hypochondria	2.00 (0.78)	—	—	—
No symptoms of hypochondria	1.71 (0.82)	—	—	—
Websites of pharmaceutical companies	—	–3.18 (294)	0.54	.002
Symptoms of hypochondria	1.67 (0.86)	—	—	—
No symptoms of hypochondria	1.21 (0.84)	—	—	—
Websites of health experts	—	–1.94 (276)	0.19	.05
Symptoms of hypochondria	2.00 (0.83)	—	—	—
No symptoms of hypochondria	1.74 (0.74)	—	—	—
Professional health portal	—	–0.67 (329)	0.12	.50
Symptoms of hypochondria	1.93 (0.67)	—	—	—
No symptoms of hypochondria	1.85 (0.69)	—	—	—
Online diagnosis system	—	–2.88 (42)	1.58	.01
Symptoms of hypochondria	2.00 (0.83)	—	—	—
No symptoms of hypochondria	0.76 (0.74)	—	—	—
Online encyclopedias	—	–2.21 (60)	0.33	.03
Symptoms of hypochondria	1.93 (72)	—	—	—
No symptoms of hypochondria	1.68 (0.80)	—	—	—
Video portals	—	–3.84 (241)	0.64	< .001
Symptoms of hypochondria	1.00 (0.87)	—	—	—
No symptoms of hypochondria	0.50 (0.67)	—	—	—

[a]Not applicable.

Evaluation of the Potential of Researched Information to Alleviate Anxiety

Overall, results showed that sharing with others affected (mean 1.94, SD 0.83) and online contact with experts (mean 1.93, SD 0.83) were able to alleviate anxiety. However, online diagnosis systems (mean 1.10, SD 0.77) and video platforms (mean 0.98, SD 0.83) seemed to cause the opposite.

However, individuals with symptoms of hypochondria did not rate health services significantly different with regard to their potential to alleviate anxiety in comparison with individuals without symptoms of hypochondria (see Table 3).

XSL•FO
RenderX

Table 3. Potential of researched information to alleviate anxiety.

Online services	Mean (SD)	*t* (*df*)	Cohen *d*	*P* value
Online contact with experts	__[a]	1.31 (25.76)	0.35	.20
Symptoms of hypochondria	1.68 (1.08)	—	—	—
No symptoms of hypochondria	2.0 (0.69)	—	—	—
Sharing with others affected	—	0.15 (254)	0.01	.88
Symptoms of hypochondria	1.95 (0.82)	—	—	—
No symptoms of hypochondria	1.93 (0.93)	—	—	—
Websites of pharmaceutical companies	—	–2.09 (167)	0.37	.04
Symptoms of hypochondria	1.64 (0.77)	—	—	—
No symptoms of hypochondria	1.37 (0.67)	—	—	—
Websites of health experts	—	–0.17 (37)	0.04	.86
Symptoms of hypochondria	1.77 (0.95)	—	—	—
No symptoms of hypochondria	1.74 (0.65)	—	—	—
Professional health portal	—	0.72 (213)	0.02	.47
Symptoms of hypochondria	1.61 (0.82)	—	—	—
No symptoms of hypochondria	1.70 (0.72)	—	—	—
Online diagnosis system	—	–0.97 (45)	0.22	.33
Symptoms of hypochondria	1.19 (0.93)	—	—	—
No symptoms of hypochondria	1.01 (0.72)	—	—	—
Online encyclopedias	—	1.59 (250)	0.27	.11
Symptoms of hypochondria	1.40 (0.80)	—	—	—
No symptoms of hypochondria	1.60 (0.71)	—	—	—
Video portals	—	0.44 (92)	0.09	.66
Symptoms of hypochondria	0.96 (0.86)	—	—	—
No symptoms of hypochondria	1.04 (0.86)	—	—	—

[a]Not applicable.

Health Behavior

Respondents seemed to be satisfied with general health care provision. Most participants were satisfied with the availability of medical services in their direct local area (mean 2.11, SD 1.07). In the last year, 80.2% (376/471) of the total sample consulted a doctor. Participants visited on average 3.55 (SD 3.01) different doctors 11.65 times (SD 16.0) in the last year.

Participants with symptoms of hypochondria considered all suggested behaviors to be more likely as a reaction to the use of online health services as participants without symptoms of hypochondria. Individuals with symptoms of hypochondria reported having visited a doctor significantly more often after health-related online searches ($P<.001$) compared with individuals without symptoms of hypochondria. Individuals with symptoms of hypochondria were also more likely to include ordering medication on the internet ($P<.001$). Furthermore, additional searches for health information as a response to the use of internet health services were considered much more probable by individuals with symptoms of hypochondria (mean 3.67, SD 1.27) than by individuals without symptoms of hypochondria (mean 2.43, SD 1.24; $P<.001$).

Discussion

Principal Findings

Cyberchondria is a form of health anxiety characterized by excessive online health research. As there have only been few elaborate studies on internet use and health fears to date, this study aimed to examine online health research of individuals with and without symptoms of hypochondria and their impact on health anxiety as well as behavior.

In the current sample of German online health service users, the portion of individuals with symptoms of hypochondria was as high as one-fourth (25.5%). In contrast, the prevalence of hypochondria in the general population is estimated at 6.7% [18]. This could be because of our recruitment methods of advertising primarily on health information websites, as persons who tend to be anxious regarding health issues consult these services more frequently [19]. Results showed that health-related internet services were used more frequently by individuals with symptoms of hypochondria. In addition, these individuals also used a greater variety of services. In line with available research, seeking online health information had a detrimental impact on

XSL•FO
RenderX

those with health anxiety [19]. Results showed that information researched online had a greater impact on health behavior in individuals with symptoms of hypochondria. Therefore, despite the established benefits of online health services such as the empowerment of patients [20], searching online health sites may carry a risk of aggravating health anxiety and lead to dysfunctional health behaviors such as doctor hopping, ordering nonprescribed medicine online, or intensifying their online research. However, studies show that the quality of information in health websites is often questionable [21]. As a consequence, individuals may have difficulties differentiating between reliable websites and websites with less reliable information. In this study, there were no significant differences in the assessment of the quality of online health services between individuals with and without symptoms of hypochondria. Similar to previous research, participants in this study perceived online health information to be somewhat reliable [22,23]. Thus, it is important to minimize misleading information on the internet. A possible solution to differentiate between websites of poor and high quality may be to label high-quality services with appropriate seals of approval and make these websites readily recognizable.

Patients suffering from hypochondriasis as well as people without this psychopathology agree that sharing worries with others—similar to an online contact with an expert—is able to alleviate health anxiety. In this sense, personal relationships, whether it is being part of a community or with a competent person, have the highest impact on gaining a feeling of security. Taking into account this relationship, an increasing intensity of medical internet research is accompanied with a greater use of medical care provision. Consequently, it is not surprising that individuals with symptoms of hypochondria reported consulting a medical professional as a consequence of online health research more often than individuals without symptoms of hypochondria. This shows that individuals with symptoms of hypochondria do not turn to online research as a result of lacking alternatives but rather consult them on- as well as offline. In conclusion, use of health-related websites may act as a catalyst for symptoms of hypochondria, reinforcing symptoms or increasing use of medical services.

Limitations and Future Research

There are some methodological limitations to this study. First, possible self-selection processes should be noted. Online surveys are predisposed for an inherent selection bias by being limited to those with access to computers and internet resources. It is possible that the number of individuals with symptoms of hypochondria in the study sample might be overestimated. Considering the significant higher amount of participants with hypochondria in this sample compared with data for the general population, this needs to be kept in mind. This might result from the fact that individuals with symptoms of hypochondria visit online health services more frequently than individuals without these symptoms [19]. In addition, to balance group sizes, participants who showed at least some symptoms of hypochondria were combined with those that could be diagnosed with hypochondria. Female participants contributed disproportionately to the respondent dataset. However, this gender bias in online surveys has been frequently observed in the literature [24]. Therefore, results cannot be considered as representative of all internet users.

To increase the range of results, a representative offline study as well as supplementary screening by the attending physician would be desirable. Thereby, self-assessment data could be validated and extended. In addition, it is necessary to investigate the association between personality traits, health-related online content, and cognitive as well as emotional processes. Thereby, studies may shift light on the role of moderating factors on the association between health anxiety and internet usage.

Conflicts of Interest

None declared.

References

1. Eichenberg C, Wolters C, Brähler E. The internet as a mental health advisor in Germany--results of a national survey. PLoS One 2013 Nov;8(11):e79206 [FREE Full text] [doi: 10.1371/journal.pone.0079206] [Medline: 24278121]
2. Starcevic V, Aboujaoude E. Cyberchondria, cyberbullying, cybersuicide, cybersex:. World Psychiatry 2015 Feb;14(1):97-100 [FREE Full text] [doi: 10.1002/wps.20195] [Medline: 25655165]
3. Huberty J, Dinkel D, Beets MW, Coleman J. Describing the use of the internet for health, physical activity, and nutrition information in pregnant women. Matern Child Health J 2013 Oct;17(8):1363-1372. [doi: 10.1007/s10995-012-1160-2] [Medline: 23090284]
4. Berle D. Cyberchondria. In: Mental Health in the Digital Age: Grave Dangers, Great Promise. Oxford: Oxford University Press; 2015:183.
5. Fergus TA. Cyberchondria and intolerance of uncertainty: examining when individuals experience health anxiety in response to Internet searches for medical information. Cyberpsychol Behav Soc Netw 2013 Oct;16(10):735-739. [doi: 10.1089/cyber.2012.0671] [Medline: 23992476]
6. Leigh H, Streltzer J. Handbook of Consultation-Liaison Psychiatry. Cham: Springer; 2015.
7. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR. Washington, DC: American Psychiatric Association; 2000.
8. Diagnostic and Statistical Manual of Mental Disorders: DSM-V. Washington, DC: American Psychiatric Association; 2013.

XSL•FO
RenderX

9. van den Heuvel OA, Veale D, Stein DJ. Hypochondriasis: considerations for ICD-11. Braz J Psychiatry 2014;36(Suppl 1):21-27 [FREE Full text] [doi: 10.1590/1516-4446-2013-1218] [Medline: 25388609]
10. White RW, Horvitz E. Cyberchondria: studies of the escalation of medical concerns in Web search. ACM Trans Inf Syst 2009 Nov 1;27(4):1-37. [doi: 10.1145/1629096.1629101]
11. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother 2013 Feb;13(2):205-213. [doi: 10.1586/ern.12.162] [Medline: 23368807]
12. McElroy E, Shevlin M. The development and initial validation of the cyberchondria severity scale (CSS). J Anxiety Disord 2014 Mar;28(2):259-265. [doi: 10.1016/j.janxdis.2013.12.007] [Medline: 24508033]
13. Bermingham S, Cohen A, Hague J, Parsonage M. The cost of somatisation among the working-age population in England for the year 2008-2009. Ment Health Fam Med 2010 Jun;7(2):71-84 [FREE Full text] [Medline: 22477925]
14. Eastin MS, Guinsler NM. Worried and wired: effects of health anxiety on information-seeking and health care utilization behaviors. Cyberpsychol Behav 2006 Aug;9(4):494-498. [doi: 10.1089/cpb.2006.9.494] [Medline: 16901253]
15. White R, Horvitz E. Web to world: predicting transitions from self-diagnosis to the pursuit of local medical assistance in web search. AMIA Annu Symp Proc 2010 Nov 13;2010:882-886 [FREE Full text] [Medline: 21347105]
16. Keller GL, Padala PR, Petty F. Clinical pearls to manage cyberchondriacs. Prim Care Companion J Clin Psychiatry 2008;10(1):75-76 [FREE Full text] [doi: 10.4088/pcc.v10n0113e] [Medline: 18311430]
17. Gmelch M, Reinecker H. [International scales for hypochondria. German-language adaptation of the Whiteley Index (WI) and the Illness Attitude Scales (IAS)]. Z Klin Psychol Psychopathol Psychother 2006 Jan;35(1):69-70. [doi: 10.1026/1616-3443.35.1.69b]
18. Bleichhardt G, Hiller W. Hypochondriasis and health anxiety in the German population. Br J Health Psychol 2007 Nov 24;12(Pt 4):511-523. [doi: 10.1348/135910706X146034] [Medline: 17931470]
19. Muse K, McManus F, Leung C, Meghreblian B, Williams JM. Cyberchondriasis: fact or fiction? A preliminary examination of the relationship between health anxiety and searching for health information on the Internet. J Anxiety Disord 2012 Jan;26(1):189-196. [doi: 10.1016/j.janxdis.2011.11.005] [Medline: 22137465]
20. van Uden-Kraan C, Drossaert C, Taal E, Seydel E, van de Laar M. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69. [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]
21. Eichenberg C, Blokus G, Malberg D. [Evidence-based patient information on the web - a study on the quality of post traumatic stress disease sites]. Z Klin Psychol Psychopathol Psychother 2013 Oct;61(4):263-271. [doi: 10.1024/1661-4747/a000169]
22. 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 US survey among 1.050 US physicians. J Med Internet Res 2003;5(3):e17 [FREE Full text] [doi: 10.2196/jmir.5.3.e17] [Medline: 14517108]
23. Peterson MW, Fretz PC. Patient use of the internet for information in a lung cancer clinic. Chest 2003 Feb;123(2):452-457. [Medline: 12576365]
24. Jackob N. [Social Research On the Internet: Methodology and Practice of the Online Survey]. Wiesbaden: Verlag für Sozialwissenschaften; 2009.

Abbreviations

DSM: Diagnostic and Statistical Manual of Mental Disorders
IAS: Illness Attitude Scale

Edited by G Eysenbach; submitted 05.05.18; peer-reviewed by S Brigitte, M Subotic-Kerry; comments to author 19.07.18; revised version received 30.09.18; accepted 27.04.19; published 10.06.19.

Please cite as:
Eichenberg C, Schott M
Use of Web-Based Health Services in Individuals With and Without Symptoms of Hypochondria: Survey Study
J Med Internet Res 2019;21(6):e10980
URL: https://www.jmir.org/2019/6/e10980/
doi:10.2196/10980
PMID:31199311

©Christiane Eichenberg, Markus Schott. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 10.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Guided Self-Help Works: Randomized Waitlist Controlled Trial of Pacifica, a Mobile App Integrating Cognitive Behavioral Therapy and Mindfulness for Stress, Anxiety, and Depression

Christine Moberg[1], PhD; Andrea Niles[2], PhD; Dale Beermann[1], MS

[1]Pacifica Labs, Inc, Minneapolis, MN, United States

[2]University of California-San Francisco, San Francisco, CA, United States

Corresponding Author:
Christine Moberg, PhD
Pacifica Labs, Inc
150 S 5th Street
Suite 825
Minneapolis, MN, 55402
United States
Phone: 1 6083470325
Email: christine.moberg@gmail.com

Abstract

Background: Despite substantial improvements in technology and the increased demand for technology-enabled behavioral health tools among consumers, little progress has been made in easing the burden of mental illness. This may be because of the inherent challenges of conducting traditional clinical trials in a rapidly evolving technology landscape.

Objective: This study sought to validate the effectiveness of Pacifica, a popular commercially available app for the self-management of mild-to-moderate stress, anxiety, and depression.

Methods: A total of 500 adults with mild-to-moderate anxiety or depression were recruited from in-app onboarding to participate in a randomized waitlist controlled trial of Pacifica. We conducted an all-virtual study, recruiting, screening, and randomizing participants through a Web-based participant portal. Study participants used the app for 1 month, with no level of use required, closely mimicking real-world app usage. Participants in the waitlist group were given access to the app after 1 month. Measurements included self-reported symptoms of stress, anxiety, depression, and self-efficacy. We performed an intent-to-treat analysis to examine the interactive effects of time and condition.

Results: We found significant interactions between time and group. Participants in the active condition demonstrated significantly greater decreases in depression, anxiety, and stress and increases in self-efficacy. Although we did not find a relationship between overall engagement with the app and symptom improvement, participants who completed relatively more thought record exercises sustained improvements in their symptoms through the 2-month follow-up to a greater degree than those who completed fewer. In addition, we found that participants who reported concomitantly taking psychiatric medications during the trial benefitted less from the app, as measured by the symptoms of anxiety and stress.

Conclusions: This study provides evidence that Pacifica, a popular commercially available self-help app, is effective in reducing self-reported symptoms of depression, anxiety, and stress, particularly among individuals who utilize thought records and are not taking psychiatric medication.

Trial Registration: ClinicalTrials.gov NCT03333707; https://clinicaltrials.gov/ct2/show/NCT03333707 (Archived by WebCite at http://www.webcitation.org/78YE07ADB)

(J Med Internet Res 2019;21(6):e12556) doi:10.2196/12556

KEYWORDS

mHealth; anxiety; depression; stress; cognitive behavioral therapy; smartphone app

XSL•FO
RenderX

Introduction

Background

In the United States, in a given year, over 19% of adults experience an anxiety disorder [1] and 6.7% experience a major depressive episode [2]. Decades of mental health care research have resulted in the development of effective nonpharmacological therapies for anxiety and depression, one of which is cognitive behavioral therapy (CBT) [3,4]. CBT is among the best-researched nonpharmacological treatments for depression and anxiety [5,6]. However, traditional CBT requires a well-trained practitioner and significant time commitment both inside and outside of the therapy room from the client; dissemination of this treatment has been somewhat limited despite its demonstrated efficacy [7].

Untreated mental illness has well-documented negative effects on physical health, health care costs, productivity, and well-being [8-10], and mental health and substance use disorders are the leading causes of disease burden in the United States [11]. Unfortunately, only 43% of people with mental illness in the United States receive professional care because of difficulties with access to care, stigma, cost, or time involved in seeking treatment [2].

Technology-enabled interventions hold great promise to ease some of the burden of mental illness due to their ability to scale and reduce barriers to entry such as cost and stigma. In fact, computerized or internet-based implementations of CBT have existed for years and many have been demonstrated to be effective for a number of mental health conditions [12-15]. However, although these products exist, they have made little impact on reducing the overall disease burden and implementation has lagged behind development [16].

Similar to the research-practice gap that exists for much of mental health treatment, digital tools that perform well in closely monitored, tightly controlled research settings do not always translate into widely utilized programs among consumers. This may be because of poor usability and design or concerns about privacy and data security [17,18]; in some cases, these apps are never made available to the public. More real-world studies that demonstrate the effectiveness of mobile-delivered treatment programs are needed.

User retention and engagement are critical to the success of digital health interventions [17,19]. This is no different than in traditional or in-person psychotherapy, where treatment retention and engagement has long been a topic of study [20]. LeBeau et al [21] defined treatment engagement as homework compliance and found that greater compliance predicted better outcomes among patients receiving CBT for anxiety disorders. However, with digital tools, engagement has often been defined by the number of uses or logins. The relationship between logins or time spent using the app or digital tool is somewhat complex; it is not always the case that greater use correlates with greater symptom improvement. For example, 1 study found that increased usage of a website for depression treatment corresponded with less symptom improvement, and the authors suggest the explanation that individuals who derived benefit from the treatment discontinued their use early [22]. Kelders et al [23] examined digital treatment adherence and identified that engagement-related analyses should consider the types of actions taken by participants (eg, didactic lessons vs feedback vs skills practice) and the time at which the individual became nonadherent (ie, early vs late). Tying engagement to outcomes can also be complex. van Gemert-Pijnen et al [24] examined participants' number of logins and use of various platform features and found a somewhat complex relationship between logins, feature use, and impact on depressive symptoms.

There is also a rich literature on predictors of treatment response in CBT, both in-person and when administered via digital tools, though it is hard to draw any clear conclusions about for whom CBT does or does not work. Høifødt et al [25] found contributions from both some demographic and clinical factors in a Web-based CBT intervention for depression. Specifically, having had more depressive episodes, being married or cohabitating, and having higher life satisfaction predicted better response. Myrh et al [26] reviewed the literature on predictors of treatment response to in-person CBT and described that although several studies have found little-to-no strong predictors of response to treatment, others have found that higher socioeconomic status and being married predict better outcomes. Dryman et al [27] conducted a study of the Internet-based CBT (ICBT) mobile app, Joyable, to examine its efficacy for treating social anxiety and found that responders (vs nonresponders) had lower baseline symptoms and spent more days in the program. Responders also called their coaches more and completed more exposures. The groups did not differ on age or gender. MoodHacker is a CBT-based app for individuals with depression. In a randomized controlled trial (RCT) examining the efficacy of MoodHacker, Birney et al [28] found that depression symptom improvements were larger among study participants who had access to an employee assistance program (EAP), though the authors note that EAP access could actually represent more than just access to an in-person counselor (eg, non-EAP participants may have different motivations than those who encountered the study through EAPs). Given these mixed data, we were also interested in examining for whom the Pacifica app would be most effective.

Objectives

Though it is not the first all-virtual RCT of a mental health support app [29,30], this study was an effort to advance the literature on the real-world use of technology-enabled CBT. We sought to test the effectiveness of Pacifica [31], a popular (over 2.4 million registered users at the time of submission) commercially available app for the self-management of stress, anxiety, and depression, in a sample of individuals with mild-to-moderate anxiety and depression who were seeking digital tools to address their mental health. We predicted that access to the app would help improve their symptoms and increase self-efficacy. In addition, we hypothesized that greater app usage, that is, more engagement, would result in larger symptom improvements. In addition, given the ongoing debate about active ingredients in therapy [32] and the transtheoretical approach of the app, we aimed to examine whether symptom improvement was related to the specific tools that were utilized (eg, thought records, meditation, and social posts). Finally, we

XSL•FO
RenderX

evaluated whether demographic variables or baseline clinical features would affect the efficacy of our app-based intervention.

Methods

Study Design

This RCT compared an immediate intervention group (*Pacifica*) with a waitlist control group (*WL*). Both groups had access to treatment as usual at their discretion. The WL group was provided access to the intervention at the end of the 1-month waiting period. The immediate intervention group was re-assessed 2 months after their completion of the 1-month intervention period to examine the stability of symptom change. The study was registered with ClinicalTrials.gov (NCT03333707) and was approved by Salus IRB (Austin, TX).

Recruitment, Screening, and Consent

We recruited participants for 3 months, from November 2017 through February 2018, via social media advertisements, a listing on ClinicalTrials.gov, and through an opt-in screen in the commercially available app itself. The vast majority of participants were recruited through the app itself. Pacifica does not advertise to acquire users, and most individuals find the app through app store searches, using terms such as *anxiety* or *stress* or via word of mouth.

We designed and built a Participant Portal that provided individuals with a *self-service* platform for screening, consent, and randomization to the app. To reduce duplicate enrollments, potential participants provided their telephone number and entered a code that was sent to them via short messaging service (SMS) text messaging. A prescreen required participants to be over the age of 18 years, fluent in English, with regular access to a smartphone, and no previous experience with Pacifica. After this prescreen, participants were provided with an informed consent document in an embedded PDF. A 4-question knowledge check confirmed that they had read and understood the document. To consent, participants entered their email address and password. Following consent, participants were screened to confirm eligibility. The Participant Portal is Web (not app) based, and therefore, participants could access the portal through any device with a Web browser. Participants assigned to the waitlist condition would not have been able to download and use the app (outside of the study) unless they signed up with a different email address than whichever address they used for this prescreen.

In addition to the criteria in the prescreen described above, the inclusion criterion was a score between 5 and 14 on the Generalized Anxiety Disorder 7-item (GAD-7) scale [33] or between 5 and 14 on the Patient Health Questionnaire 8-item (PHQ-8) scale [34]. Exclusion criteria were (1) score below 5 on GAD-7 and PHQ-8 or above 14 on either, (2) positive response on screener for previous diagnosis of bipolar disorder, schizophrenia, other psychotic spectrum disorder, or organic brain disease, and (3) currently pregnant. Given the all-virtual nature of the study, these criteria were selected to match the intensity of the intervention to the severity of the users, allowing us to balance participant safety with external validity and value of the data.

An initial group of 9279 individuals completed the prescreen, 1524 completed informed consent, and 500 were randomized to groups. In addition, 7 individuals' participation was discontinued for failing to download the app within 48 hours of screening and randomization. Of the 253 individuals who were assigned to the treatment group, 182 participants used Apple (iOS) devices and 61 used Android devices. Furthermore, 3 participants accessed Pacifica with both iOS and Android devices. Participants were assessed using questionnaire measures, the details for which are below. See the CONSORT flow diagram (Figure 1) for reasons for exclusion and experimental compliance.

Intervention

Participants in the Pacifica group were given access to Pacifica Premium, versions 5.7 through 5.9.1. Updates to the app that occurred during the research study were bug fixes and performance improvements and would not have impacted the therapeutic approach or usability of the app. *Descriptions of the app below are as it was in the research study and may not be exactly the same as the currently available product.* Pacifica is a mobile app marketed as a guided self-help tool for the management of stress, anxiety, and depression. It is not described as a treatment for any particular diagnosis nor is it described as a substitute for professional treatment. At onboarding, users select up to 3 goals on which to work from a list of 8 options. The app prompts users once per day to rate their mood and, based on their mood rating, recommends activities to improve their mood via *Suggested Activities* (ie, ecological momentary intervention; see the studies by Schueller et al and Lovibond and Lovibond [35,36]). There are also 35 days of *Guided Paths* that are approximately 10-min audio psychoeducational lessons with paired activities. Alternatively, users can use the app *buffet style*, using whichever tools they find helpful, whenever they choose. Participants in the study were not provided guidance or suggestions regarding the amount of level of app usage beyond any prompts in the app (described above).

Descriptions and details on the different activities are given below:

Mood

Allows users to rate their mood on a scale from *Great* to *Awful*. Users may optionally label specific emotions or attach a journal-style note to their mood rating.

Health

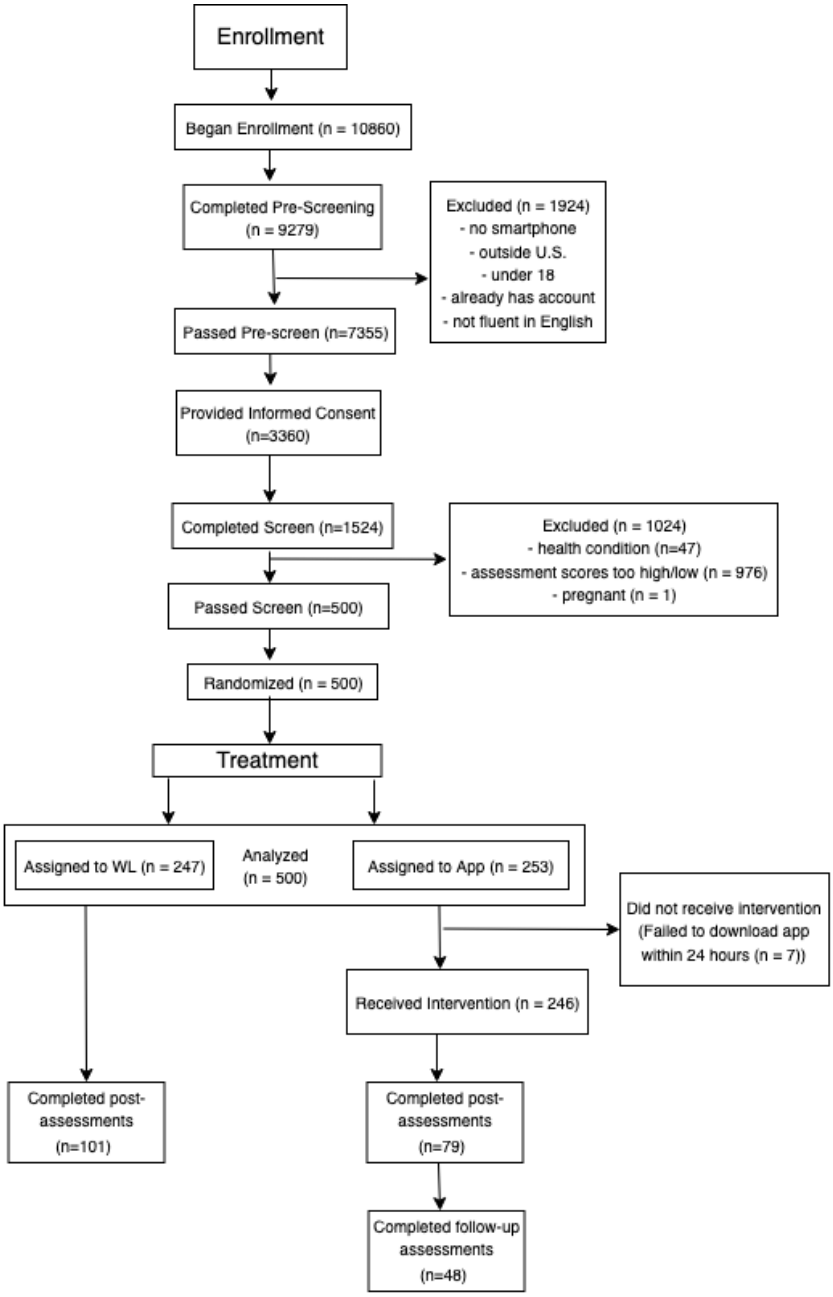
Allows users to track health behaviors such as sleep, caffeine consumption, and exercise. Users can customize their health behaviors and goals for each health behavior (ie, how many hours of sleep or number of cups of coffee). They may optionally set up an alarm so that Pacifica sends them a notification to remind them to enter their health data.

Meditation/Relax

Offers over 40 audio activities, most of which are intended to promote mindfulness or relaxation. These include deep breathing, progressive muscle relaxation, and a variety of mindfulness activities.

XSL•FO
RenderX

Figure 1. Consolidated Standards of Reporting Trials diagram.



Thoughts

Offers 9 different activities to help users re-examine their thoughts, identify cognitive distortions, and reframe thoughts. The prompts are divided into 3 groups: *basics*, which help the user recognize the relationship between their thoughts and emotions, learn about cognitive distortions, and reframe negative thoughts; *journal prompts*, which allow for free writing; and *advanced tools*. The advanced tools include one that creates a pie chart to help users identify other factors beyond themselves that may contribute to events, a tool that asks a series of questions to help them evaluate evidence, and an activity that guides them to drill down and identify core beliefs. All of these tools are guided exercises and do not provide automatic analysis of the individual's text inputs.

Goals

Allows the user to create a list of challenges to complete to meet their longer-term goals. A fairly generic list tool, it can be customized to address anxiety by enabling the user to build and address a hierarchy of feared or avoided situations. Users coping with depression can use the list to engage in behavioral activation to re-engage with life. The tool prompts the user to rate the perceived difficulty of the item before and after engagement to help them recognize that their initial estimation of the challenge may not be accurate. There are example items that may be chosen or users can type in their own goals.

Guided Paths

Offer 35 days of psychoeducational content to teach the user how to utilize the tools in the app and help maintain motivation and interest. The lessons also provide background information about CBT and mindfulness. Each 5- to 10-min audio lesson is

XSL•FO
RenderX

paired with an in-app activity. The audio content is presented either as a didactic session or a mock session between a client a therapist. The 35 days of content are divided into 4 paths: one which offers an introduction to each of the tools in the app, 2 which are sequential and offer a deeper focus on CBT, and one which focuses on building mindfulness skills.

Hope

On the basis of the principles of distress coping, the hope board is a user's personal repository of inspirational quotes and images. It also allows the user to save their completed *challenges* to help them feel a sense of accomplishment and savor their wins.

Community

A peer-support community that is not moderated by professionals, but is rather a place where users can post their thoughts, challenges, and questions and receive support from others using the app. The community is anonymous but utilizes a flagging system to remove disruptive or inappropriate posts or users. Users are provided the rules and guidelines of the community when they first access the community.

Progress

Allows the user to review their mood ratings and completed activities. The tool graphs their mood ratings against health behaviors and other in-app activities so that users can identify patterns and triggers. There is also a *skills* tab that visually displays the user's completed activities, separated by type.

Emergency Resources

Includes listing of crisis lines and resources for users in emergency situations.

Measurement

Questionnaire Measures

Study participants were assessed at 3 time points: at baseline (pretreatment), 4 weeks later (ie, posttreatment; at the end of the intervention period), and 3 months after baseline/2 months postintervention (follow-up; only the Pacifica group completed the follow-up because the WL group was allowed to access the intervention after the 4-week assessment). At each assessment, participants were both emailed and notified via SMS that they had available assessments to complete. Assessments were completed via the secure Participant Portal described above (ie, not in the app itself).

The Depression Anxiety and Stress Scales-21 (DASS-21) [37] is a widely used freely available self-report measure of stress, anxiety, and depression. A shortened version of the DASS-42, its 21 items are measured on a Likert scale and yield 3 subscale scores: stress, anxiety, and depression. The DASS-21 consists of statements such as "I couldn't seem to experience any positive feeling at all" with 4 response options ranging from "Did not apply to me at all" to "Applied to me very much or most of the time." For each subscale, summed scores range from 0 to 21, with higher scores indicating greater symptoms. Its internal consistency and concurrent validity have been shown to be in the acceptable-to-excellent range [38].

The PHQ-8 [34] is an 8-item self-report measure of depression. It is identical to the popular PHQ-9 [39] measure but eliminates the question that queries suicidal ideation; it was selected given the all-virtual nature of the study. An example item from the PHQ-8 is "Little interest or pleasure in doing things," with response options ranging from "Not at all" to "Nearly every day." Scores on the PHQ-8 range from 0 to 24, and the measure has been shown to have good internal consistency and reliability [34]. Higher scores indicate greater symptoms.

The GAD-7 [33] is a widely used, freely available, 7-item self-report measure to address severity of Generalized Anxiety Disorder. An example item is "Feeling nervous, anxious, or on edge," with response options ranging from "Not at all" to "Nearly every day." Summed scores on the GAD-7 range from 0 to 21, with higher scores indicating greater symptoms. The GAD-7 has good reliability and validity [33].

The General Self-Efficacy Scale [40] is a 10-item self-report measure that asks participants about their beliefs regarding their ability to cope with stressors or challenges in their life. Example items include "I can always manage to solve difficult problems if I try hard enough," with response options ranging from "Not at all true" to "Exactly true." Scores on the General Self-Efficacy Scale range from 10 to 40, with higher scores indicating greater self-efficacy. The measure has been shown to be reliable and valid [41].

Overall, 2 composites were created from the 2 measures of depression (PHQ-8 and DASS-21 Depression Subscale) and the 2 measures of anxiety (GAD-7 and DASS-21 Anxiety Subscale) to generate more reliable and valid indices of symptoms. We standardized each measure at the pretreatment time point, then used the means and SDs calculated at pretreatment to standardize these measures at subsequent timepoints. After the scores were standardized, measures were averaged to create a single measure of each construct. These measures are referred to as *depression composite* and *anxiety composite* in the results section and have a mean of zero and a SD of one 1 pretreatment.

Process Measures

To better understand possible mediators of symptom improvement and app efficacy, we examined participants' number of logins and number of completed activities as measures of app engagement. It should be noted that for the purposes of this study, *activities* include *thoughts*, *goals* (either setting or completing), *relax*, and *community* (writing a post). The engagement analysis did not include mood or health ratings, *hope* posts, or time spent reviewing content in the community forums because of the complexity involved in data extraction, cleaning, and formatting for these activities.

Moderator Measures

To examine whether baseline patient characteristics predicted treatment response to Pacifica versus WL, we tested demographic and clinical characteristics. Demographic features included age (treated as a continuous variable), gender (0=female, 1=male), income (8 levels from <$20K to >$200K treated as a continuous variable), marital status (0=not married, 1=married), education (0=did not complete 4-year college,

XSL•FO
RenderX

1=completed 4-year college), and race (0=other, 1=white). Clinical features included current (at screening) use of psychiatric medications (0=no, 1=yes), current (at screening) use of psychotherapy (0=no, 1=yes), use of other mental health apps (0=no, 1=yes), and whether they reported being diagnosed with depression (0=no, 1=yes) or anxiety (0=no, 1=yes).

Statistical Analyses

Power Analysis

On the basis of a two-week pilot study among college undergraduates, we anticipated an effect size of .3 (Cohen *d*). We targeted 90% power to detect an effect to a significance level of .05 in a 2-group 1-tailed *t* test, which revealed a sample of 191 participants per group. The previous study had about 75% completion, so we targeted to randomize 250 participants per group. Though the use of pilot studies for estimating effect sizes is questionable (cf [42]), this pilot study was the most relevant data available on which to base our power analysis. The effect sizes found in this study are larger than those in the pilot study: between .4 and .54 for the key symptoms measures versus .3 in the pilot.

We were not able to meet our target of 191 completers because of greater-than-anticipated attrition: 101 WL participants and 79 active group participants completed the study. Given this number of completers, we were able to detect an effect size of .38 with 80% power given a significance level of .05 and a 2-group, 2-tailed *t* test.

Statistical Analysis

Data were analyzed using multilevel modeling (MLM) in Stata 15.0 [43]. MLM accounts for nesting of time points within subjects, which allows for examination of change within and between subjects across time (pretreatment, posttreatment, and follow-up) and by group (WL and Pacifica). MLM includes all participants with at least 1 measurement, and thus, it can be used to estimate intent-to-treat effects. Separate models were run for each of the 4 dependent variables (depression composite, anxiety composite, stress, and general self-efficacy). Statistical significance was defined as falling below a threshold of alpha=.05, as this is the standard threshold used in randomized clinical trials in psychology.

To examine the main effects of the Pacifica intervention compared with WL, time was modeled at level 1, with 1 segment specified between the pre- and posttreatment time points, and a second segment between the post and follow-up time points. This approach models typical trends in treatment studies where the greatest effects occur by posttreatment and changes level-off over follow-up and has been used in a number of previous studies [44,45,46,47]. Group (WL vs Pacifica) was modeled at level 2. Models included the main effects of the 2 time segments, the main effect of group, and the interaction between group and time. We examined the interaction between group and time for statistical significance. All models included random effects of the intercept and time, which were determined to be significant based on likelihood ratio tests. Between-group differences were assessed using the *margins* command in Stata by comparing the slopes from pre- to posttreatment and by comparing the predicted means from the model at the posttreatment time point. Effect sizes are calculated based on the method described by Feingold [48] that produces estimates analogous to Cohen *d* for growth curve models in randomized clinical trials. As the WL group received the intervention after the posttreatment time point, we were unable to compare the groups at follow-up. Thus, for the follow-up time point, the slope of change from posttreatment to follow-up was tested only for the Pacifica group.

To examine whether engagement with the app was related to treatment response, we tested whether overall engagement (number of app logins) and specific types of engagement (number of times using the goals, relaxation, community, and thoughts sections of the app), were associated with symptom change in the Pacifica group. Models included the main effect of engagement, the main effects of the 2 time segments, and the interactions between engagement and time. We examined the interactions between the engagement variables and each time segment for statistical significance. Thus, we examined whether engagement predicted change from pre- to posttreatment and from posttreatment to follow-up for each type of engagement and for each dependent variable.

To examine moderators of response to Pacifica versus WL, we tested whether demographic features (age, gender, income, marital status, educational level, and race) and clinical features (use of psychiatric medications, engagement in psychotherapy, use of other mental health apps, and previously diagnosed anxiety and depression) were associated with symptom change from pre- to posttreatment in the Pacifica group compared with the WL group. Models included the main effects of the moderator, the main effect of time, the main effect of group, all 2-way interactions, and the 3-way interaction. We examined the 3-way interaction between the moderator, group, and time for statistical significance. Significant 3-way interactions were followed up with tests of simple interaction effects of group by time at different levels of the moderator.

Results

Pretreatment Group Differences

Table 1 presents demographic data on the WL and Pacifica groups. Randomization was successful, and the 2 groups were very similar with regard to demographics; we found no significant differences between the groups (p s>.210). The sample was largely female (75%) and white (80%) and had at least some college education (91%).

Study Attrition

Of the 500 participants randomized, 204 completed both pre- and posttreatment assessments. This represents 47% of WL participants and 35% of Pacifica participants. The difference in attrition between the groups was significant: χ_1^2 (n=500)=7.7; P=.006. We examined the baseline symptom ratings for the individuals who dropped out versus those who did not and found no differences (see Table 2). Further, no differences were found when these comparisons were stratified by group (P>.18).

XSL•FO
RenderX

Table 1. Descriptive data of sample.

Characteristic	Waitlist control group (n=247)	Active (n=253)	Chi-square (*df* ${3.1}	*P* value
Age (years), mean (SD)[a]	30.2 (10.8)	30.2 (10.9)	—[b]	.96
Gender, n (%)			**2.73 (4)**	**.60**
Male	56 (23)	54 (21)	—	—
Female	84 (74)	190 (75)	—	—
Transgender man	4 (2)	2 (1)	—	—
Nonconforming	3 (1)	6 (2)	—	—
Race, n (%)			**3.73 (5)**	**.59**
Asian	11 (4)	10 (4)	—	—
Black	20 (9)	28 (11)	—	—
Hawaii/Pacific Islander	1 (0.5)	2 (1)	—	—
Native American/Alaska Native	5 (2)	5 (2)	—	—
White	208 (84)	202 (80)	—	—
No reply	2 (1)	6 (2)	—	—
Education level, n (%)			**2.80 (5)**	**.73**
Some high school	1 (0.5)	4 (2)	—	—
High School	21 (9)	16 (6)	—	—
Some college	81 (33)	86 (34)	—	—
Associates	19 (8)	17 (7)	—	—
Bachelor’s	80 (32)	84 (33)	—	—
Master’s or higher	44 (17)	46 (18)	—	—
No reply	1 (0.5)	0	—	—
Marital Status, n (%)			**1.50 (4)**	**.83**
Never married	161 (65)	161 (64)	—	—
Married	63 (26)	67 (26)	—	—
Separated	3 (1)	6 (2)	—	—
Divorced	19 (8)	17 (7)	—	—
Widowed	1 (0.5)	2 (1)	—	—
Income level in USD, n (%)			**9.62 (7)**	**.21**
Under 20K	101 (41)	97 (38)	—	—
20-35k	50 (20)	39 (15)	—	—
35-50k	38 (15)	38 (15)	—	—
50-75k	25 (10)	46 (18)	—	—
75-100k	17 (7)	20 (8)	—	—
100-150k	9 (4)	9 (4)	—	—
150-200k	3 (1)	3 (1)	—	—
Over 200k	4 (2)	1 (0)	—	—
Using other apps for mental health			**0.09 (1)**	**.77**
Yes	52 (21)	56 (22)	—	—
No	195 (79)	197 (78)	—	—
Currently in therapy			**0.47 (1)**	**.49**
Yes	57 (23)	52 (21)	—	—
No	190 (77)	201 (79)	—	—

XSL•FO
RenderX

Characteristic	Waitlist control group (n=247)	Active (n=253)	Chi-square (*df* ${3.1}	*P* value
Currently taking psychiatric medication			**1.61 (2)**	**.45**
Yes	91 (37)	101 (40)	—	—
No	155 (6)	151 (60)	—	—
No response	1 (0)	1 (0)	—	—
Previous diagnosis of depression			**0.13 (1)**	**.72**
Yes	184	192	—	—
No	63	54	—	—
Previous diagnosis of anxiety			**0.13 (1)**	**.72**
Yes	205	213	—	—
No	42	40	—	—

[a]t_{498}=0.05.

[b]Not applicable.

Table 2. Mean values on baseline outcomes for participants who dropped versus completed the posttreatment assessment.

Measure	Drop	Complete	*t* (*df*)	P value
Depression composite	−0.01	0.01	0.31 (498)	.76
Anxiety composite	0.04	−0.05	−1.16 (498)	.25
Depression Anxiety and Stress Scales-21 stress	8.77	8.55	−0.78 (498)	.44
General self-efficacy	26.97	26.75	−0.57 (498)	.57

Key Outcome: Symptom Change

We first examined whether, compared with individuals in the WL group, individuals assigned to use Pacifica experienced greater improvement in anxiety, depression, stress, and self-efficacy. Our intent-to-treat analysis revealed significant group × time interactions for each of the outcome measures (depression composite, anxiety composite, DASS-21 stress, and self-efficacy) such that change in the Pacifica group was greater than change in the WL group. These results were unchanged when individual measures of depression and anxiety were examined separately. See Multimedia Appendix 1 for means and SEs of the individual depression and anxiety measures. In addition, to assess whether treatment dropout affected these results, we analyzed additional models that included the group × dropout interaction and main effects of group and dropout, and the results were unchanged. Table 3 provides parameter estimates for the group × time interaction from pre to posttreatment, 95% CIs, significance levels, and effect sizes for each outcome measure.

We additionally examined the simple effect of time within groups. Table 4 shows that the Pacifica group experienced significant change from pre to post for each variable (decreases for depression, anxiety, and stress and an increase for self-efficacy). Anxiety level in the WL group significantly decreased from pre to post. For the Pacifica group, from post to follow-up, there were no significant changes in any outcome. There were significant group differences at post for each measure (see Figure 2). The WL group was higher than the Pacifica group for depression symptoms (0.66; CI 0.37 to 0.94; *P*<.001), anxiety symptoms (0.47; CI 0.17 to 0.77; *P*=.002), and DASS-21 stress (2.09; CI 1.09 to 3.08; *P*<.001), and lower on self-efficacy (−1.46; CI −2.62 to −0.30; *P*=.014).

Table 3. Group × time interaction.

Measure	Beta	95% CI	*P* value	Effect size (*d*)
Depression composite	−0.59	−0.86 to -.03	<.001	0.54
Anxiety composite	−0.43	−0.71 to −0.15	.003	0.40
Depression Anxiety and Stress Scales-21 stress	−1.79	−2.74 to −0.84	<.001	0.46
General self-efficacy	1.55	0.53 to 2.58	.003	0.34

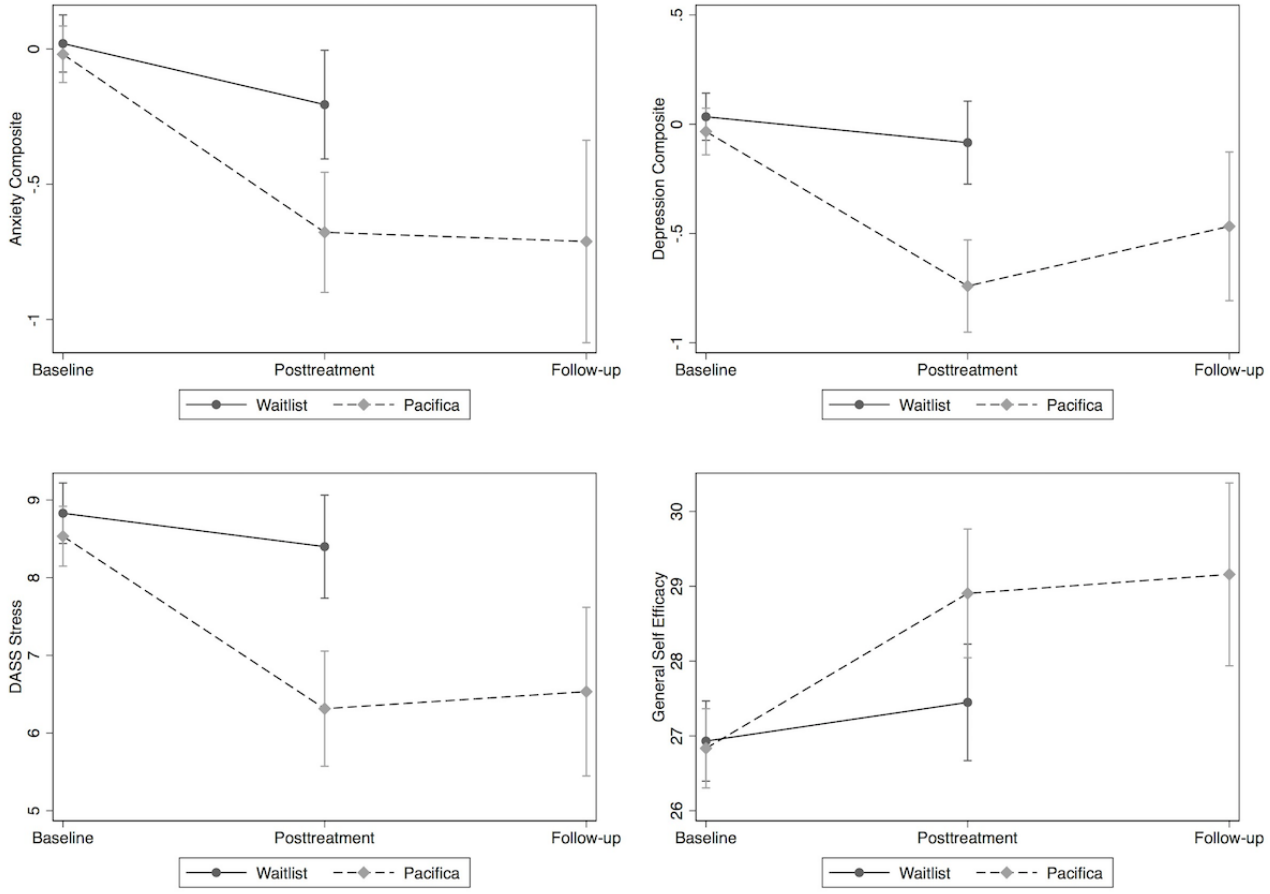
XSL•FO
RenderX

Table 4. Score change on key measures across time by group.

Measure	Change across time	95% CI	*P* value
Depression composite			
Pre versus post			
Control	−0.12	−0.30 to 0.06	.21
Pacifica	−0.71	−0.91 to 0.50	<.001
Post versus Follow-up			
Pacifica	0.27	−0.10 to 0.64	.149
Anxiety composite			
Pre versus post			
Control	−0.23	−0.41 to −0.04	.018
Pacifica	−0.66	−0.87 to −0.45	<.001
Post versus Follow-up			
Pacifica	−0.03	-0.44 to 0.37	.87
Depression Anxiety and Stress Scales-21 stress			
Pre versus post			
Control	−0.43	−1.06 to 0.20	.18
Pacifica	−2.22	−2.93 to −1.51	<.001
Post versus Follow-up			
Pacifica	0.22	−0.97 to 1.41	.72
General self-efficacy			
Pre versus post			
Control	0.52	−0.16 to 1.2	.13
Pacifica	2.10	1.30 to 2.84	<.001
Post versus Follow-up			
Pacifica	0.25	−1.02 to 1.53	.697

XSL•FO
RenderX

Figure 2. Key outcome measures by group and time. DASS: Depression Anxiety Stress Scales.



Clinically Significant Change

Participants were classified as having achieved clinically significant change (CSC) on the PHQ-8 and GAD-7 as outlined by Jacobson and Truax [49]. The individual scales were used for the CSC analysis because test-retest reliability metrics were needed to calculate the reliable change index. Rates of CSC at post for PHQ-8 were as follows: Pacifica, 41.8% (n=33/79); WL, 16.8% (n=17/101); χ_1^2 (N=180)=13.74; *P*<.001. Rates of CSC at post for GAD-7 were as follows: Pacifica, 39.2% (n=31/79); WL, 24.2% (n=24/99); χ_1^2 (N=178)=4.63; *P*=.031. The rate of CSC at follow-up for the Pacifica group was 35.4% for the PHQ-8 and 48.9% for the GAD-7.

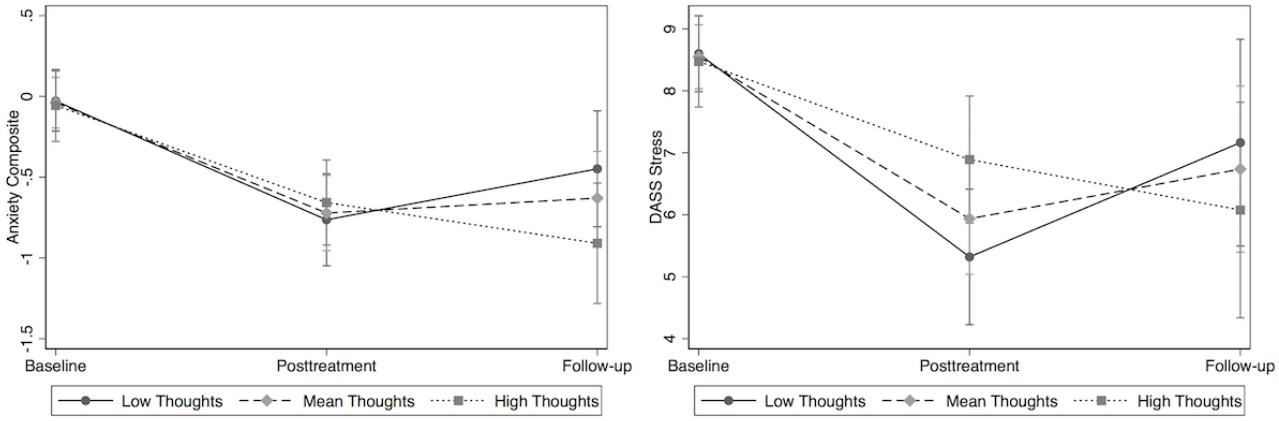
Engagement Effects

The median number of logins in the Pacifica group during the 30-day intervention was 19, with a range of 1 to 286. We did not find any significant interactions for group by time by *overall* engagement with the app (defined by total number of logins) for depression composite (*P*>.21), anxiety composite (*P*>.34), stress (*P*>.11), or self-efficacy (*P*>.55).

However, for tests examining unique activities in the app, we found that the use of thought record tools was significantly associated with symptom improvement for 2 of the outcome measures: anxiety composite and stress (see Figure 3). Specifically, for the anxiety composite, greater number of thought records completed during treatment was associated with greater anxiety reduction from posttreatment to follow-up (beta =-.10; CI −0.19 to −0.02; *P*=.019), although not from pretreatment to posttreatment (*P*=.406) . Number of thought records completed was not significantly associated with baseline anxiety in the MLM (*P*=.834). For DASS-21 stress, a greater number of thought records completed during treatment was associated with less stress reduction from pre- to posttreatment (beta = −.30; CI −.50 to −.11; *P*<.01), but greater stress reduction from posttreatment to follow-up (beta =-.47; CI −.87 to −.08; *P*<.05). Number of thought records completed was not significantly associated with baseline stress in the MLM (*P*=.772). Descriptively, individuals who completed relatively more thought records showed a delayed positive effect in terms of anxiety and stress reduction. Use of the thought record tool was not significantly associated with change in depression (*P*>.13) or general self-efficacy (*P*>.48). The use of the goals, relax, and community tools was not associated with change in anxiety (*P*>.18), depression (*P*>.13), stress (*P*>.29), or general self-efficacy (*P*>.26).

XSL•FO
RenderX

Figure 3. Interaction between use of "Thoughts" tool and anxiety and stress outcomes. DASS: Depression Anxiety Stress Scales.



Moderator Analysis

Demographics

We examined the 3-way interaction between group, time, and demographic responses to examine whether any of these demographic traits moderated treatment response.

Age, gender, income, marital status, educational level, and race did not significantly moderate the effect of group on symptom improvement. See Table 5 for details.

Clinical Features

We also conducted moderator analyses for each of the clinical characteristics listed in Table 1. We found only 1 characteristic that significantly interacted with group and time to predict outcome measures: presence/absence of concomitant psychiatric medications at baseline. As shown in Figure 4, the presence of psychiatric medications moderated the group effect such that the benefits of Pacifica compared with WL were significantly greater for people who were not taking psychiatric medications than for people who were taking medications. This moderation of the group × time interaction only reached significance for the anxiety composite and DASS-21 stress measures.

For the anxiety composite, tests of simple effects showed that for participants not taking psychiatric medications, those assigned to Pacifica had a significantly greater reduction in anxiety than those assigned to WL (*slope difference*=−.68, CI −1.01 to −0.35; *P*<.001). However, for those taking psychiatric medications, no difference in symptom reduction emerged between Pacifica and WL (*P*=.771). A similar pattern emerged for stress. For participants not taking psychiatric medications, those assigned to Pacifica had a significantly greater reduction in stress than those assigned to WL (*slope difference*=−2.54; CI −3.70 to −1.39]; *P*<.001). However, for those taking medications, no difference in symptom reduction emerged between Pacifica and WL (*P*=.615).

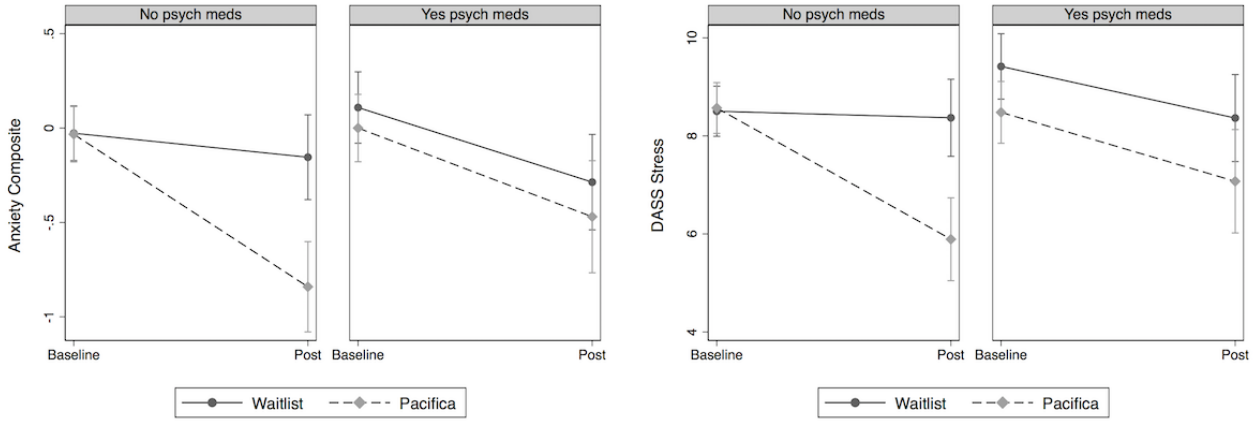
Table 5. Group × time × moderator.

Variable	Depression, beta (95% CI)	Anxiety, beta (95% CI)	DASS-21 stress, beta (95% CI)	Self-efficacy, beta (95% CI)
Age (years)	−.01 (−0.03 to 0.02)	−.01 (−0.03 to 0.02)	.01 (−0.08 to 0.09)	−.03 (−0.13 to 0.06)
Gender	−.16 (−0.81 to 0.49)	−.08 (−0.73 to 0.57]	−1.72 (−4.02 to 0.57)	−.90 (−3.40 to 1.61)
White	−.22 (−0.93 to 0.48)	−.40 (−1.11 to 0.30)	.18 (−2.29 to 2.65)	.50 (−2.20 to 3.20)
College educated	.15 (−0.36 to 0.65)	.48 (−0.03 to 0.99)	.60 (−1.20 to 2.39)	1.35 (−0.63 to 3.34)
Married	−.25 (−0.81 to 0.32)	−.45 (−1.02 to 0.12)	−1.85 (−3.85 to 0.15)	1.30 (−0.88 to 3.49)
Income level	.04 (−0.12 to 0.19)	.09 (−0.07 to 0.25)	.27 (−0.29 to 0.84)	−.08 (−0.70 to 0.53)
Using other apps	.01 (−0.59 to 0.62)	−.18 (−0.79 to 0.42)	−.31 (−2.44 to 1.81)	.67 (−1.68 to 3.02)
Receiving psychotherapy	.17 (−0.46 to 0.81)	.44 (−0.20 to 1.08)	.52 (−1.76 to 2.80)	−1.81 (−4.39 to 0.78)
Taking psychiatric medication	.47 (−0.04 to 0.98)	.61 (0.09 to 1.12)[a]	2.19 (0.39 to 3.99)[a]	.18 (−1.84 to 2.20)
Diagnosed with depression	−.27 (−0.84 to 0.30)	.13 (−0.44 to 0.70)	.30 (−1.72 to 2.32)	.80 (−1.40 to 3.00)
Diagnosed with anxiety	.21 (−0.51 to 0.93)	−.02 (−0.75 to 0.72)	.17 (−2.38 to 2.73)	−1.18 (−4.00 to 1.64)

[a]Effects are significant at *P*<.05.

XSL•FO
RenderX

Figure 4. Moderation of group by time effect by participant self-reported psychiatric medication use. DASS: Depression Anxiety Stress Scales.



Discussion

This study is the first published RCT of Pacifica, a widely available, popular smartphone app–based intervention for self-help of mild-to-moderate stress, anxiety, and depression. The tools implemented in the intervention are based on the integration of CBT, mindfulness, and mood and health tracking. Results indicated that this intervention is effective compared with a waitlist control at 1 month in reducing self-reported symptoms of depression, anxiety, and stress and increasing feelings of self-efficacy. Between-group effect sizes were in the small-to-medium range. In addition, treatment gains were maintained for 2 months following the end of the 1-month intervention period, especially for anxiety symptoms. These results are particularly noteworthy given the very light touch nature of the study, short duration of the intervention period, and the real-world sample of individuals.

App usage was self-guided, and participants were not instructed to use the app with any specific level of frequency, so rather than consider adherence, we explored whether amount of usage (ie, number of completed activities) was associated with symptom improvement. We did not find any association with overall usage, possibly because looking at the total number of app logins is a simplistic way of quantifying engagement. However, we did find that individuals who completed relatively more thought records demonstrated delayed improvement, which is noteworthy. Though there have been challenges in the literature to the importance of cognitive restructuring in CBT [50], this study suggests that, at least for technology-enabled CBT, thought records were helpful. Subjectively, completing thought records is a more time-intensive and cognitively demanding activity than health tracking or meditation and increased use of these activities may be a marker of greater commitment to treatment or improvement. We should note, though, that in this study thought records included both basic journaling prompts and more complex reframing activities. However, this study was unfortunately underpowered to examine the effects of different types of thought tools. The apparent delayed positive benefit of completing relatively more thought activities is consistent with the common perception that although examining and challenging negative thoughts is difficult in the short term, it can result in more long-lasting benefits. In addition, there are data suggesting that CBT is effective at preventing relapse to anxiety and depression [51,52], but further research is needed to more completely understand these effects, particularly in computerized CBT.

Demographic characteristics among participants were not related to outcomes, which is consistent with the mixed findings across previous studies examining traditional (ie, not online delivered) CBT [53-55] as well as internet-based CBT [13,25,56-60]. This suggests that app-based CBT can have effects across diverse populations rather than being relevant only for a specific group. However, we found differential effects between participants who were and were not taking psychiatric medications at baseline. Though not significant for measures of depression or self-efficacy, for the anxiety composite and DASS-21 stress measures, the presence of medication at baseline diminished the difference between the groups, that is, both the WL and app groups improved. Inspection of the means reveals not that the presence of medications markedly reduced the effectiveness of the app, but rather that there was little added benefit of the app above and beyond the medication.

Although this finding is inconsistent with some data that combined treatments perform better than therapy alone [61], it is consistent with literature that has found minimal benefit of combining medication with psychotherapy [62]. Certain medications such as benzodiazepines may interfere with behavioral therapy [63]. To be sure, though, a guided self-help tool such as Pacifica is not psychotherapy—nor does it purport to be. The personal relationship established with a practitioner may be what catalyzes treatment and nudges it above medication alone; blended models that incorporate both self-help solutions with human coaching have been found to be superior to self-help alone [64]. The drawback of these models is their ability to scale. Stepped care models that triage services and match the level of in-person support with client need are likely the way of the future [65]. Such models combine the rapport and human connection of a human therapist or coach with the ubiquity and outside-of-the-clinic support of a mobile app. The question of whether combining mobile interventions with pharmacotherapy is effective is particularly relevant given the fact that a majority of individuals receive their psychotropic medications from primary care providers [66] who typically only see them

XSL•FO
RenderX

intermittently; digital tools could potentially provide additional support between appointments.

This project represents a step forward in that it is a real-world evaluation of a popular, commercially available technology-enabled intervention. However, there are several limitations that should be noted. First, the study was conducted among a convenience sample that was rather homogeneous in its demographic makeup. The participants were largely college-educated white females. Previous research has found that women are overrepresented in Web-based research studies [67]. Although this sample may represent the individuals who are most likely to utilize technology-enabled mental health interventions, this homogeneity may limit the degree to which the outcomes can be generalized. Owing to the fact that no demographic information is collected from users upon sign-up, we cannot assess how representative this sample is of all of the individuals who utilize Pacifica.

Although the participant pool was homogeneous from a demographic perspective, the study was conducted in a sample that may have been heterogeneous with respect to diagnosis. We did not conduct diagnostic interviews and only have participants' responses to symptom rating questionnaires. In the real world, many individuals seeking Web-based tools to manage their mental health are doing so outside of the context of the therapy clinic and may not have formal diagnoses. It is noteworthy that our sample had relatively low levels of anxiety and depression. This further highlights the fact that individuals who seek digital tools may be on the mild end of the spectrum and that these tools may need to be used in conjunction with therapy to be effective for more severely affected individuals. Future research should examine whether these data can be replicated in a more diverse sample to consider whether efficacy varies based on demographics. It would also be important to verify the app's efficacy among individuals who have been formally diagnosed versus those who self-selected and were simply screened using self-reported measures. Furthermore, analyses of engagement with the app were exploratory, and significant findings should be replicated in future studies.

Another limitation relates to the finding regarding the moderation of the effects by the use of psychotropic medications. Although this finding is interesting, because we had minimal information about participants' use of medications (ie, no data on the types, dosage, or any change to medication across the course of the study), this should be examined further in future work.

This study provides encouraging findings around the ability of popular, commercially available, guided self-help tools to empower individuals to manage their symptoms of stress, anxiety, and depression and increase their self-efficacy. This study also serves as a proof point for an all-virtual study. Given the limitations of existing health care systems and the obstacles to care that exist for individuals coping with mental health conditions, mobile apps and technology-enabled interventions can play an important role in expanding access and serving as an adjunct to in-person treatment. Future research should continue to clarify the best application of these tools and how they can be better integrated into existing workflows and care delivery systems.

Acknowledgments

This study was funded by Pacifica Labs, Inc.

Conflicts of Interest

At the time of this study, CM was the Head of Psychology, owned options in Pacifica Labs, Inc, and was receiving a salary from the company. DB was the Cofounder and CEO of Pacifica Labs, Inc, owned a large share of the company's stock, and received a salary from the company.

Multimedia Appendix 1

Predicted means and standard errors for individual depression (Patient Health Questionnaire, Depression Anxiety Stress Scales-Depression) and anxiety (Generalized Anxiety Disorder 7-item, Depression Anxiety Stress Scales-Anxiety) scales.

[PDF File (Adobe PDF File), 15KB - jmir_v21i6e12556_app1.pdf]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - jmir_v21i5e12556_fig.pdf]

References

1. National Comorbidity Survey (NCS). 2017. URL: https://www.hcp.med.harvard.edu/ncs/index.php [accessed 2019-05-20] [WebCite Cache ID 78WbZE7mg]
2. Key Substance Use and Mental Health Indicators in the United States: Results from the 2016 National Survey on Drug Use and Health. Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration; 2017.

XSL•FO
RenderX

3. Beck J. Cognitive Behavior Therapy, Second Edition. New York, NY: Guilford Press; 2011:9781609185046.
4. DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, et al. Cognitive therapy vs medications in the treatment of moderate to severe depression. Arch Gen Psychiatry 2005 Apr;62(4):409-416. [doi: 10.1001/archpsyc.62.4.409] [Medline: 15809408]
5. Butler A, Chapman J, Forman E, Beck A. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. Clin Psychol Rev 2006 Jan;26(1):17-31. [doi: 10.1016/j.cpr.2005.07.003] [Medline: 16199119]
6. Hofmann S, Asnaani A, Vonk I, Sawyer A, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
7. Bruns E, Kerns S, Pullmann M, Hensley S, Lutterman T, Hoagwood KE. Research, data, and evidence-based treatment use in state behavioral health systems, 2001-2012. Psychiatr Serv 2016 Dec 1;67(5):496-503 [FREE Full text] [doi: 10.1176/appi.ps.201500014] [Medline: 26695495]
8. Insel T. Assessing the economic costs of serious mental illness. Am J Psychiatry 2008 Jun;165(6):663-665. [doi: 10.1176/appi.ajp.2008.08030366] [Medline: 18519528]
9. Keyes CG, Grzywacz JG. Health as a complete state: the added value in work performance and healthcare costs. J Occup Environ Med 2005 May;47(5):523-532. [Medline: 15891532]
10. Rice D, Miller LS. Health economics and cost implications of anxiety and other mental disorders in the United States. Br J Psychiatry Suppl 1998(34):4-9. [Medline: 9829010]
11. Kaiser Family Foundation Analysis of Data from Institute for Health MetricsEvaluation. Peterson-Kaiser Health System Tracker. 2015. Global Burden of Disease Study, 2015 URL: https://www.healthsystemtracker.org/chart-collection/current-costs-outcomes-related-mental-health-substance-abuse-disorders/ [accessed 2019-05-20] [WebCite Cache ID 78WcPqAMz]
12. Espie C, Kyle S, Williams C, Ong J, Douglas N, Hames P, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep 2012 Jun 1;35(6):769-781 [FREE Full text] [doi: 10.5665/sleep.1872] [Medline: 22654196]
13. Hedman E, Furmark T, Carlbring P, Ljótsson B, Rück C, Lindefors N, et al. A 5-Year follow-up of internet-based cognitive behavior therapy for social anxiety disorder. J Med Internet Res 2011 Jun 15;13(2):e39 [FREE Full text] [doi: 10.2196/jmir.1776] [Medline: 21676694]
14. Kaltenthaler E, Parry G, Beverley C, Ferriter M. Computerised cognitive-behavioural therapy for depression: systematic review. Br J Psychiatry 2008 Sep;193(3):181-184. [doi: 10.1192/bjp.bp.106.025981] [Medline: 18757972]
15. Pretorius N, Arcelus J, Beecham J, Dawson H, Doherty F, Eisler I, et al. Cognitive-behavioural therapy for adolescents with bulimic symptomatology: the acceptability and effectiveness of internet-based delivery. Behav Res Ther 2009 Sep;47(9):729-736. [doi: 10.1016/j.brat.2009.05.006] [Medline: 19515360]
16. Vis C, Mol M, Kleiboer A, Bührmann L, Finch T, Smit J, et al. Improving implementation of eMental health for mood disorders in routine practice: systematic review of barriers and facilitating factors. JMIR Ment Health 2018 Mar 16;5(1):e20 [FREE Full text] [doi: 10.2196/mental.9769] [Medline: 29549072]
17. Torous J, Nicholas J, Larsen M, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
18. Stawarz K, Preist C, Tallon D, Wiles N, Coyle D. User experience of cognitive behavioral therapy apps for depression: an analysis of app functionality and user reviews. J Med Internet Res 2018 Dec 6;20(6):e10120 [FREE Full text] [doi: 10.2196/10120] [Medline: 29875087]
19. Alkhaldi G, Hamilton F, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 8;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
20. Baekeland F, Lundwall L. Dropping out of treatment: a critical review. Psychol Bull 1975 Sep;82(5):738-783. [Medline: 1103201]
21. Lebeau RT, Davies C, Culver N, Craske M. Homework compliance counts in cognitive-behavioral therapy. Cogn Behav Ther 2013;42(3):171-179. [doi: 10.1080/16506073.2013.763286] [Medline: 23419077]
22. Clarke G, Kelleher C, Hornbrook M, Debar L, Dickerson J, Gullion C. Randomized effectiveness trial of an Internet, pure self-help, cognitive behavioral intervention for depressive symptoms in young adults. Cogn Behav Ther 2009;38(4):222-234 [FREE Full text] [doi: 10.1080/16506070802675353] [Medline: 19440896]
23. Kelders SM, Bohlmeijer ET, van Gemert-Pijnen JW. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013 Aug 20;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
24. van Gemert-Pijnen JE, Kelders SM, Bohlmeijer ET. Understanding the usage of content in a mental health intervention for depression: an analysis of log data. J Med Internet Res 2014 Jan 31;16(1):e27 [FREE Full text] [doi: 10.2196/jmir.2991] [Medline: 24486914]

XSL•FO
RenderX

25. Høifødt RS, Mittner M, Lillevoll K, Katla S, Kolstrup N, Eisemann M, et al. Predictors of response to web-based cognitive behavioral therapy with high-intensity face-to-face therapist guidance for depression: a Bayesian analysis. J Med Internet Res 2015 Sep 2;17(9):e197 [FREE Full text] [doi: 10.2196/jmir.4351] [Medline: 26333818]
26. Myhr G, Talbot J, Annable L, Pinard G. Suitability for short-term cognitive-behavioral therapy. J Cogn Psychother 2007 Dec 1;21(4):334-345. [doi: 10.1891/088983907782638743]
27. Dryman M, McTeague L, Olino T, Heimberg R. Evaluation of an open-access CBT-based Internet program for social anxiety: patterns of use, retention, and outcomes. J Consult Clin Psychol 2017 Oct;85(10):988-999. [doi: 10.1037/ccp0000232] [Medline: 28650193]
28. Birney A, Gunn R, Russell J, Ary D. MoodHacker mobile web app with email for adults to self-manage mild-to-moderate depression: randomized controlled trial. JMIR Mhealth Uhealth 2016 Jan 26;4(1):e8 [FREE Full text] [doi: 10.2196/mhealth.4231] [Medline: 26813737]
29. Anguera J, Jordan J, 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]
30. Arean P, Hallgren K, Jordan J, Gazzaley A, Atkins D, Heagerty P, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
31. Pacifica L. Pacifica: Version 5.7 through 5.9.1. San Francisco, CA: Pacifica Labs, Inc; 2017.
32. Wampold B. How important are the common factors in psychotherapy? An update. World Psychiatr 2015 Oct;14(3):270-277 [FREE Full text] [doi: 10.1002/wps.20238] [Medline: 26407772]
33. Spitzer R, Kroenke K, Williams J, 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]
34. Kroenke K, Strine T, Spitzer R, Williams J, Berry J, Mokdad A. 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]
35. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
36. Schueller S, Aguilera A, Mohr D. Ecological momentary interventions for depression and anxiety. Depress Anxiety 2017 Dec;34(6):540-545. [doi: 10.1002/da.22649] [Medline: 28494123]
37. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety Stress Scales. Second Edition. Sydney, Aus: Psychology Foundation; 1995:733414230.
38. Antony M, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychol Asses 1998 Jun;10(2):176-181. [doi: 10.1037/1040-3590.10.2.176]
39. Kroenke K, Spitzer R, Williams J. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
40. Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs. Windsor, England: NFER-NELSON; 1995:35-37.
41. Scholz U, Gutiérrez Doña B, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. Eur J of Psych Assess 2002 Sep;18(3):242-251. [doi: 10.1027//1015-5759.18.3.242]
42. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
43. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp; 2017.
44. Craske M, Niles A, Burklund L, Wolitzky-Taylor K, Vilardaga J, Arch J, et al. Randomized controlled trial of cognitive behavioral therapy and acceptance and commitment therapy for social phobia: outcomes and moderators. J Consult Clin Psychol 2014 Dec;82(6):1034-1048 [FREE Full text] [doi: 10.1037/a0037212] [Medline: 24999670]
45. Niles A, Craske M, Lieberman M, Hur C. Affect labeling enhances exposure effectiveness for public speaking anxiety. Behav Res Ther 2015 May;68:27-36. [doi: 10.1016/j.brat.2015.03.004] [Medline: 25795524]
46. Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, et al. Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. J Am Med Assoc 2010 May 19;303(19):1921-1928 [FREE Full text] [doi: 10.1001/jama.2010.608] [Medline: 20483968]
47. Wolitzky-Taylor K, Niles A, Ries R, Krull J, Rawson R, Roy-Byrne P, et al. Who needs more than standard care? Treatment moderators in a randomized clinical trial comparing addiction treatment alone to addiction treatment plus anxiety disorder treatment for comorbid anxiety and substance use disorders. Behav Res Ther 2018 Dec;107:1-9. [doi: 10.1016/j.brat.2018.05.005] [Medline: 29800622]
48. Feingold A. Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. Psychol Methods 2009 Mar;14(1):43-53 [FREE Full text] [doi: 10.1037/a0014699] [Medline: 19271847]
49. Jacobson N, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [Medline: 2002127]

XSL•FO
RenderX

50. Longmore RJ, Worrell M. Do we need to challenge thoughts in cognitive behavior therapy? Clin Psychol Rev 2007 Mar;27(2):173-187. [doi: 10.1016/j.cpr.2006.08.001] [Medline: 17157970]
51. DiMauro J, Domingues J, Fernandez G, Tolin DF. Long-term effectiveness of CBT for anxiety disorders in an adult outpatient clinic sample: a follow-up study. Behav Res Ther 2013 Feb;51(2):82-86. [doi: 10.1016/j.brat.2012.10.003] [Medline: 23262115]
52. Hollon S, DeRubeis R, Shelton R, Amsterdam J, Salomon R, O'Reardon J, et al. Prevention of relapse following cognitive therapy vs medications in moderate to severe depression. Arch Gen Psychiatry 2005 Apr;62(4):417-422. [doi: 10.1001/archpsyc.62.4.417] [Medline: 15809409]
53. Button K, Turner N, Campbell J, Kessler D, Kuyken W, Lewis G, et al. Moderators of response to cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care. J Affect Disord 2015 Mar 15;174:272-280. [doi: 10.1016/j.jad.2014.11.057] [Medline: 25527998]
54. Fournier J, DeRubeis R, Shelton R, Hollon S, Amsterdam J, Gallop R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. J Consult Clin Psychol 2009 Aug;77(4):775-787 [FREE Full text] [doi: 10.1037/a0015401] [Medline: 19634969]
55. Wolitzky-Taylor KB, Arch JJ, Rosenfield D, Craske MG. Moderators and non-specific predictors of treatment outcome for anxiety disorders: a comparison of cognitive behavioral therapy to acceptance and commitment therapy. J Consult Clin Psychol 2012 Oct;80(5):786-799. [doi: 10.1037/a0029418] [Medline: 22823858]
56. El Alaoui S, Ljótsson B, Hedman E, Kaldo V, Andersson E, Rück C, et al. Predictors of symptomatic change and adherence in internet-based cognitive behaviour therapy for social anxiety disorder in routine psychiatric care. PLoS One 2015;10(4):e0124258 [FREE Full text] [doi: 10.1371/journal.pone.0124258] [Medline: 25893687]
57. Berger T, Hämmerli K, Gubser N, Andersson G, Caspar F. Internet-based treatment of depression: a randomized controlled trial comparing guided with unguided self-help. Cogn Behav Ther 2011 Dec;40(4):251-266. [doi: 10.1080/16506073.2011.616531] [Medline: 22060248]
58. Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, Vermeulen H, et al. Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cogn Behav Ther 2009 Dec;38(4):206-221. [doi: 10.1080/16506070802408086] [Medline: 19221919]
59. Spek V, Cuijpers P, Nyklícek I, Smits N, Riper H, Keyzer J, et al. One-year follow-up results of a randomized controlled clinical trial on internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years. Psychol Med 2008 May;38(5):635-639. [doi: 10.1017/S0033291707002590] [Medline: 18205965]
60. Vernmark K, Lenndin J, Bjärehed J, Carlsson M, Karlsson J, Oberg J, et al. Internet administered guided self-help versus individualized e-mail therapy: a randomized trial of two versions of CBT for major depression. Behav Res Ther 2010 May;48(5):368-376. [doi: 10.1016/j.brat.2010.01.005] [Medline: 20152960]
61. Cuijpers P, Sijbrandij M, Koole S, Andersson G, Beekman A, Reynolds C. Adding psychotherapy to antidepressant medication in depression and anxiety disorders: a meta-analysis. World Psychiatry 2014 Feb;13(1):56-67 [FREE Full text] [doi: 10.1002/wps.20089] [Medline: 24497254]
62. Otto M, Smits J, Reese H. Combined psychotherapy and pharmacotherapy for mood and anxiety disorders in adults: Review and analysis. Clin Psych Sci Pract 2005;12:72-86. [doi: 10.1093/clipsy.bpi009]
63. Otto MW, McHugh R, Kantak KM. Combined pharmacotherapy and cognitive-behavioral therapy for anxiety disorders: medication effects, glucocorticoids, and attenuated treatment outcomes. Clin Psychol (New York) 2010 Jun 1;17(2):91-103 [FREE Full text] [doi: 10.1111/j.1468-2850.2010.01198.x] [Medline: 26855480]
64. 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]
65. Scogin F, Hanson A, Welsh D. Self-administered treatment in stepped-care models of depression treatment. J Clin Psychol 2003 Mar;59(3):341-349. [doi: 10.1002/jclp.10133] [Medline: 12579549]
66. Olfson M, Blanco C, Wang S, Laje G, Correll C. National trends in the mental health care of children, adolescents, and adults by office-based physicians. JAMA Psychiatry 2014 Jan;71(1):81-90. [doi: 10.1001/jamapsychiatry.2013.3074] [Medline: 24285382]
67. Karyotaki E, Riper H, Twisk J, Hoogendoorn A, Kleiboer A, Mira A, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 1;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]

Abbreviations

CBT: cognitive behavioral therapy
CSC: clinically significant change
DASS-21: Depression Anxiety Stress Scales-21
EAP: employee assistance program
GAD-7: Generalized Anxiety Disorder 7-item
MLM: multilevel modeling

XSL•FO
RenderX

PHQ-8: Patient Health Questionnaire 8-item
RCT: randomized controlled trial
SMS: short messaging service

Edited by G Eysenbach; submitted 19.10.18; peer-reviewed by S Schueller, N Karnik, E Lattie; comments to author 20.12.18; revised version received 11.04.19; accepted 28.04.19; published 08.06.19.

Please cite as:
Moberg C, Niles A, Beermann D
Guided Self-Help Works: Randomized Waitlist Controlled Trial of Pacifica, a Mobile App Integrating Cognitive Behavioral Therapy and Mindfulness for Stress, Anxiety, and Depression
J Med Internet Res 2019;21(6):e12556
URL: https://www.jmir.org/2019/6/e12556/
doi:10.2196/12556
PMID:31199319

©Christine Moberg, Andrea Niles, Dale Beermann. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding Long-Term Trajectories in Web-Based Happiness Interventions: Secondary Analysis From Two Web-Based Randomized Trials

Christopher A Sanders[1], MA; Stephen M Schueller[2], PhD; Acacia C Parks[3], PhD; Ryan T Howell[4], PhD

[1]Department of Psychological Sciences, University of Missouri, Columbia, Columbia, MO, United States

[2]Department of Psychological Science, University of California Irvine, Irvine, CA, United States

[3]Happify Health, New York, NY, United States

[4]Psychology Department, San Francisco State University, San Francisco, CA, United States

Corresponding Author:
Christopher A Sanders, MA
Department of Psychological Sciences
University of Missouri, Columbia
210 McAlester Hall
320 S 6th St
Columbia, MO, 65211
United States
Phone: 1 573 882 6860
Email: c.a.sanders@mail.missouri.edu

Abstract

Background: A critical issue in understanding the benefits of Web-based interventions is the lack of information on the sustainability of those benefits. Sustainability in studies is often determined using group-level analyses that might obscure our understanding of who actually sustains change. Person-centric methods might provide a deeper knowledge of whether benefits are sustained and who tends to sustain those benefits.

Objective: The aim of this study was to conduct a person-centric analysis of longitudinal outcomes, examining well-being in participants over the first 3 months following a Web-based happiness intervention. We predicted we would find distinct trajectories in people's pattern of response over time. We also sought to identify what aspects of the intervention and the individual predicted an individual's well-being trajectory.

Methods: Data were gathered from 2 large studies of Web-based happiness interventions: one in which participants were randomly assigned to 1 of 14 possible 1-week activities (N=912) and another wherein participants were randomly assigned to complete 0, 2, 4, or 6 weeks of activities (N=1318). We performed a variation of *K*-means cluster analysis on trajectories of life satisfaction (LS) and affect balance (AB). After clusters were identified, we used exploratory analyses of variance and logistic regression models to analyze groups and compare predictors of group membership.

Results: Cluster analysis produced similar cluster solutions for each sample. In both cases, participant trajectories in LS and AB fell into 1 of 4 distinct groups. These groups were as follows: those with high and static levels of happiness (n=118, or 42.8%, in Sample 1; n=306, or 52.8%, in Sample 2), those who experienced a lasting improvement (n=74, or 26.8% in Sample 1; n=104, or 18.0%, in Sample 2), those who experienced a temporary improvement but returned to baseline (n=37, or 13.4%, in Sample 1; n=82, or 14.2%, in Sample 2), and those with other trajectories (n=47, or 17.0%, in Sample 1; n=87, or 15.0% in Sample 2). The prevalence of depression symptoms predicted membership in 1 of the latter 3 groups. Higher usage and greater adherence predicted sustained rather than temporary benefits.

Conclusions: We revealed a few common patterns of change among those completing Web-based happiness interventions. A noteworthy finding was that many individuals began quite happy and maintained those levels. We failed to identify evidence that the benefit of any particular activity or group of activities was more sustainable than any others. We did find, however, that the distressed portion of participants was more likely to achieve a lasting benefit if they continued to practice, and adhere to, their assigned Web-based happiness intervention.

(J Med Internet Res 2019;21(6):e13253) doi:10.2196/13253

XSL•FO
RenderX

KEYWORDS

cluster analysis; depression; happiness; random allocation

Introduction

Background

There is a wide variety of electronic health (eHealth) and mobile health (mHealth) interventions available to anyone who is interested in Web-based mental health care or self-improvement and has access to the internet or a mobile network. The individual goals of these Web-based interventions range from smoking cessation [1] to the prevention of weight gain [2] to even the treatment of posttraumatic stress disorder and depression [3]. Many interventions, however, lack a specific focus and instead attempt to build general wellness or *happiness*. These interventions have been called elsewhere online positive psychological interventions (OPPIs) [4]. OPPIs have been the subject of much research and development, given people's strong interest in pursuing and increasing happiness.

Although the efficacy of many of these Web-based interventions, including OPPIs, has been demonstrated in the peer-reviewed academic literature, little information is available on the sustainability of the benefits that people see when they use these interventions. Although it has been suggested that such information might be useful to allow users to make educated choices about eHealth and mHealth interventions [5], the methods to produce such an understanding have not been well developed [6]. In a review of 11 OPPI efficacy studies featuring randomized controlled trial designs, 4 studies (approximately 36%) did not report effects beyond the posttest [4]. The remaining 7 studies included in the review, however, provide initial evidence that the increases to happiness following an OPPI can be partially sustained for 6 weeks [7], 3 months [8], and 6 months [9,10] after the completion of the intervention and that a remediation of depression symptoms can similarly be observed 3 months [11] and 6 months [9,10,12,13] after exposure to an OPPI.

Although these studies do provide initial support for the idea that the benefits of an OPPI can be sustained, OPPI efficacy trials with randomized designs and some degree of follow-up assessment are limited in number beyond this subset [4] and, to the best of our knowledge, none of them have conducted specific analyses to understand the longitudinal outcomes. To address this discrepancy, in this analysis, we apply a person-centric or *idiographic* approach to post-OPPI follow-up assessments to identify which outcome trajectories are most likely before exploring individual differences in distinct longitudinal outcome trajectories.

Online Positive Psychological Interventions and the People Who Use Them

OPPIs are most often brief, skill-based exercises that are intended to improve happiness and well-being by teaching individuals the cognitive and behavioral strategies of chronically happy people [14]. These interventions are technological translations drawing from the broader area of positive psychological interventions (PPIs) which have generally demonstrated efficacy in using a variety of strategies and delivery modalities [15,16]. Bolier and Abello [4] reviewed the evidence specifically for OPPIs and not only found that effects tended to be smaller than those found generally for offline PPIs but also noted that the only direct comparison at the time of publication between a Web-based and offline intervention found no significant difference between them [17]. Thus, it is likely that many of the benefits accrued from PPIs also apply to OPPIs.

Meta-analyses have revealed that most studies, be it Web-based or computer-based PPIs, have evaluated only the immediate impact of the intervention [15,16], comparing baseline reports of happiness with assessments of happiness made immediately following the intervention. We mentioned earlier that a smaller subset of studies have continued to collect data for 3 or more months after the completion of a given intervention [7,9-11,13,18], and these studies tend to find evidence, at the level of group-wise analyses, that many happiness interventions continue to benefit individuals who practice them through this period. This line of research, however, leaves a number of open questions regarding the nuances of longitudinal outcomes, particularly regarding individual differences.

Any person-centric approach should also appreciate the types of people who tend to seek out and use OPPIs. Fortunately, the characteristics and behaviors of such people have been investigated in past research. Parks et al [19] provide a thoughtful analysis of the type of lay people that free OPPIs appeal to, a group that the authors refer to as Web-based happiness seekers. These authors draw 3 main conclusions. First, they found that roughly half of the seekers are somewhat happy people intending to achieve greater happiness, but that the other half are quite distressed and some might even be experiencing a mental health condition. Second, they found that, overall, happiness seekers tend to frequently employ several activities in their pursuit of happiness and may persist with these activities for several months. Third, when happiness seekers are provided with easy access to a variety of (presumably) happiness-promoting activities, it is the frequency and variety of activities that they engage in that predict increases to mood and happiness.

It is also worth noting, however, that Web-based happiness seekers might be more motivated than others to increase their happiness or report their happiness as increasing. One investigation of such people found that even those exposed to hypothesized-inert psychoeducational material received boosts in happiness and well-being [20]. It is worth noting that the characteristics of Web-based happiness seekers may not be that different from those of people seeking other forms of psychological interventions on the Web and that the type of engagement with OPPIs is consistent with other eHealth and mHealth interventions in which they require a substantial degree of self-motivation and self-guidance. As such, the understanding of OPPIs and Web-based happiness seekers can help contribute to an understanding of eHealth and mHealth interventions more generally.

XSL•FO
RenderX

Other studies have explored those who self-select into PPIs, although not completely in a Web-based environment. For example, Kaczmarek et al [21] allowed college students to, voluntarily and anonymously, self-initiate a Web-based gratitude intervention after they completed a separate study. The 11.5% of participants who started the Web-based happiness intervention were more likely than their peers to express high levels of trait curiosity and endorse strong intentions to change their lifestyle. However, although Parks et al [19] found that the prevalence of depression symptoms was higher in Web-based happiness seekers than that found in the general population, Kaczmarek et al [21] found that depressive symptoms were related to a reduced tendency to start the intervention. Lyubomirsky et al [18] also conducted a similar study on positive interventions in which participants self-selected into either a study advertised as consisting of cognitive exercises or a study advertised as consisting of happiness exercises, with all participants randomly assigned to receive either a positive intervention or control exercise. In this study, they found no initial differences between the conditions on well-being, which would seem to be more in line with the findings of Parks et al [19] than those of Kaczmarek et al [21]. They did, however, find that the only people to significantly experience an increase in well-being after the intervention were those who sought out a happiness exercise in the first place and were administered a PPI (rather than a control exercise). This supports the finding of Haeck et al [20], in which it suggests that those who are motivated and interested in happiness-increasing activities might be more successful potentially because of increased effort or motivation. Bearing this in mind, it is worth examining the characteristics and behaviors of those who experience long-term benefits from PPIs, and specifically OPPIs, to better understand the mechanisms underlying the benefits.

Hedonic Adaptation

One challenge in evaluating the long-term benefits of OPPIs is that happiness naturally fluctuates to some extent over time [22]. Furthermore, if OPPIs are to be lastingly effective, then they must overcome a psychological homeostatic process called hedonic adaptation [23,24]. Hedonic adaptation is the process through which most people revert to a previous and stable level of well-being even after significant life events and changes [25]. The earliest study noting this phenomenon was done on rare events, such as winning the lottery or having a limb amputated (ie, Brickman et al [26]), but a robust body of evidence also demonstrates this phenomenon in more common, major life events including job change (eg, was found by Chadi & Hetschko [27]), childbirth (eg, was found by Dyrdal & Lucas [28]), and marital divorce (see Kramrei et al [29] for a meta-analysis of divorce effects). In all these cases, the changes to happiness (whether positive or negative) that people saw following these events were generally temporary and usually dissipated completely within a few months or a year.

Hedonic adaptation is likely due to a combination of affective, cognitive, and motivational processes [30,31]. In the Sustainable Happiness Model, Lyubomirsky et al [32] argue that because one's genetic set point accounts for around 50% of the variance in happiness and circumstances account for only about 10% of the variance, up to 40% of the variance in individual happiness is because of intentional activities, that is, things people think and do. These intentional activities can be teachable and are ultimately the focus of OPPIs. Sheldon et al [33] propose the Hedonic Adaptation Prevention model to suggest that hedonic adaptation is not inevitable but instead can be counteracted by intentional activity. Specifically, the Hedonic Adaptation Prevention model suggests that appreciation, surprise, variety, and intrinsic change are all mechanisms through which sustained benefits in happiness are achievable. Lyubomirsky and Layous [34] extend some of the thinking in this model by suggesting that the characteristics of both the intervention and the person contribute to the ultimate benefit received. In addition to aspects, such as variety, they also highlight issues such as dosage, social support, and triggers of the intervention as well as baseline affective states of the person as being drivers of happiness change. Their study guides this research by identifying the characteristics and behaviors that might be worth exploring in terms of mechanisms and long-term OPPI benefits.

Primary Analyses of the Current Datasets

The data analyzed here originate from 2 large samples of Web-based happiness seekers, both of which have been reported elsewhere [19,35]. Although previous reports on these data have been restricted to the baseline characteristics of participants, immediate outcomes, and attrition, this paper reports the long-term outcomes and attempts to understand who achieves sustainable long-term benefits. The authors of the report on the first sample in our paper [19] conducted cluster analysis using baseline reports of depression symptoms, life satisfaction (LS), and affect balance (AB) and found that the participants fell into 1 of 2 groups. About half (49.5%) fell into a *distressed* group reporting low initial levels of well-being and high levels of depressive symptoms, whereas the other half (50.5%) formed a *nondistressed* cluster reporting high initial levels of well-being and low levels of depressive symptoms. Noteworthy, the *distressed* cluster had an average level of depressive symptoms of mean 26.74 (SD 10.58) on the Center for Epidemiological Studies–Depression Scale (CES-D; [36]), whereas the *nondistressed* cluster had an average level of depressive symptoms of mean 7.93 (SD 5.85) on the same scale. The authors argue that Web-based happiness seekers can be reasonably categorized into those who might be suffering from current mental health problems and those who are not and that the success of a positive intervention could be largely dependent on what category a participant belongs to.

Another report on these data was published by 2 of the authors of this paper (SMS and ACP) [35] who conducted a 6-week OPPI wherein the participants were assigned to 2, 4, or 6 weeks of intervention content or an assessment-only control condition. Participants received a new intervention each week such that the 2-week condition comprised 2 different PPIs, the 4-week condition comprised 4 different PPIs, and the 6-week condition comprised 6 different PPIs. The 2- and 4-week interventions were more effective at reducing symptoms than the control condition or the 6-week intervention by the end of the study period. Although participants in the 6-week condition did not obtain gains beyond those seen in the 4-week condition, those in the 6-week condition were more likely than other groups to

XSL•FO
RenderX

continue practicing some of the exercises. In their interpretation of these mixed findings, the authors [35] explained that

> *"It might be that increasing the diversity of exercises leads to participants splitting their time among the techniques and not focusing on any of the techniques long enough to benefit substantially."*

Both previous papers on these data, however, were also limited by their use of nomothetic techniques to focus on the overall benefits that people obtained rather than trying to understand the trajectories of individual benefit. Although the first set of authors [19] did create clusters of *distressed* and *nondistressed* participants, this clustering was created at baseline and did not explore how information could be gained about individual differences in change trajectories.

This Analysis

In a secondary analysis of these 2 large datasets (described by [19] and [35]), we explore longitudinal outcomes in 2 separable dimensions of happiness (LS and mood) after self-selected exposure to OPPIs. To best understand how happiness change occurs and is maintained, we adopt a person-centric approach that identifies the trajectories in happiness reports.

By examining LS and AB across the months following self-selection into a Web-based happiness intervention, we sought to address the following questions:

1. Do people typically experience adaptation after these interventions and return to baseline levels of well-being after the OPPI or are they more likely to achieve a sustainable well-being increase during this period?
2. If adaptation is prevalent yet avoidable, what can people do to prevent it?
3. How many other distinct well-being trajectories occur during this timeframe and what are their shapes?

The analysis we present here addresses these questions by identifying the most common outcome trajectories that people have reported following exposure to an OPPI, grouping people based on those trajectories and using group membership as a proxy to explore individual differences in longitudinal outcomes.

Methods

Recruitment

In both the samples, participants were directed to a common research portal via Web-based advertisements to participate in a research study on positive psychology exercises and a printed advertisement in Seligman's *Authentic Happiness* (2002) book. No compensation, beyond the advertised benefit of participating in a happiness-boosting intervention, was offered to the participants. This study was approved by the Institutional Review Board (IRB) at the University of Pennsylvania under an exempt IRB as a process of continuous quality improvement, thus it was not deemed to be a clinical trial and was not registered as such.

Participants

The 2 samples differed in the period during which data were collected. Those who enrolled between July 2006 and February 2007 appeared in Sample 1, whereas those who enrolled between February 2007 and November 2008 were represented in Sample 2.

The 2230 participants across Sample 1 (n=912) and Sample 2 (n=1318) were, demographically, very similar. Both samples contained more women than men (76.5% overall), and the 2 samples did not seem to differ in this regard (z=.496, P=.617) and were made up of people who were moderately educated (74.3% overall had a Bachelor's degree; 70% in Sample 1 and 77.2% in Sample 2; z=-3.853; P<.001) and middle-aged (mean age 43.5 years overall; 45.3 years in Sample 1 and 42.3 years in Sample 2; t_{2296}=5.944; P<.001). As described by Parks et al [19], these individuals were, in addition, likely to be distinct from other research samples because they represent the self-selected individuals who actively seek to increase their happiness through Web-based interventions and other mediums. Thus, the results of this paper were specific to this unique population: a group that includes many highly distressed persons seeking to overcome their depression symptoms.

Procedures

On enrollment, participants provided consent and answered demographic questions along with a set of surveys relevant to their mental health and well-being. Afterward, they were randomly assigned to Web-based happiness intervention conditions that differed between samples.

Sample 1: Individual Interventions

Participants in Sample 1 were randomly assigned between conditions in a 14-group randomized controlled trial design. A total of 13 of these conditions represent hypothesized 1-week happiness activities (eg, writing a gratitude letter or savoring a beautiful day), whereas the final condition was based on a 1-week active control writing task used in previous studies (see [9]). Intervention instructions varied considerably by condition and are summarized in Multimedia Appendix 1.

Sample 2: Multiple Interventions

Participants in Sample 2 were randomly assigned between 4 conditions in which they received 0, 2, 4, or 6 weeks of positive psychology exercises, with a new exercise administered within each week of their assigned intervention. The 0 exercise condition was included as a waitlist control condition and participants could receive exercises after 6 weeks of completing assessments only. The exercises included in the experimental conditions were a subset of those to which Sample 1 was randomly assigned. Exercises were provided in a fixed order such that people in the 2-, 4-, or 6-week conditions in Sample 2 received the same content across conditions (with the variance in content being due to a variance in the intervention length). The administration order for the exercises is also provided in Multimedia Appendix 1 alongside the description of each exercise.

Measures

Across the 2 samples, the surveys administered overlapped considerably. As such, the measures presented here were used in both the samples except where specified. Participants in both the samples completed a battery of surveys before intervention

XSL•FO
RenderX

assignment, at the end of each week during the intervention period, and 1, 3, 6, and 12 months after the intervention period. These survey batteries contained the measures we described here, as well as others. Unfortunately, item-level data were not available for the calculation of scale reliability statistics in these samples. This occurred because scales were summarized by the survey software before data export and the item-level data used to create these summaries were no longer available.

Life Satisfaction

The Satisfaction with Life scale [37] is a global measure of LS that consists of 5 items, each with a 7 point Likert scale response (ranging from 1= *strongly disagree* to 7= *strongly agree*). Higher scores on each of the items indicate greater LS. The scale itself has demonstrated strong internal consistency, test-retest reliability, and construct validity in numerous studies (see [38] for a review). The items on this scale are relatively face-valid: "I am satisfied with my life." and "If I could live my life over, I would change almost nothing." are 2 examples. The metric used in this paper to assess LS is the sum of an individual's scores on the items in this scale, with higher numbers reflecting greater satisfaction.

Affect Balance

The Positive and Negative Affect schedule [39] is a 20-item scale that asks participants to rate the extent to which they are currently experiencing each of the 20 different emotions on a 5 point Likert scale (ranging from 1=*very likely or not at all* to 5=*extremely*). A total of 10 items corresponded to positive emotions (eg, *enthusiastic* and *proud*), whereas the other 10 items reflected negative emotions (eg, *upset* and *nervous*). AB was calculated by summing the scores across positive items and subtracting the sum score across negative items.

Depression Symptoms

The CES-D scale [36] is a commonly used metric for current depression symptoms in research settings that asks participants to report the frequency with which they had experienced 20 different symptoms (eg, restlessness, loneliness, and loss of appetite) over the past week. Although the CES-D is not designed for the specific diagnosis of clinical depression, it has repeatedly demonstrated good reliability and construct validity across a wide variety of clinical [40] and nonclinical samples (see [41] for review) as a tool for identifying subclinical depression symptoms. The 20 items on this scale are each accompanied by a 4 point Likert scale response (ranging from 0=*rarely or none of the time* to 3=*most or almost all the time*) and these responses are summed to provide a proxy of depression symptom prevalence.

Use and Adherence

Participants in both samples were asked at each assessment after the posttest to report the number of days over the past week in which each assigned positive psychology exercise was performed (responses ranged from 0 to 7 days) and whether the specific instructions were adhered to (participants indicated *Yes* or *No* for each assigned exercise). As the number of exercises that were assigned to each participant varied within and between samples, responses to these 2 items were summed within timepoint for each participant and grouped based on the number of assigned activities before being z-transformed within each group.

Statistical Analysis

We adopted the following 3-step analytic plan with a person-centric focus so that we might address our research questions with the greatest accuracy. We excluded all participants who completed fewer than 4 assessments of either LS or AB. The remaining missing data were deleted on a pairwise basis.

1. Trace a trajectory in well-being over time for each individual participant
2. Group similar trajectories based on their shape and identify which curve shapes are most common
3. Interpret the most common trajectory shapes and evaluate their relation to depression symptomology and use statistics

In the first step of our process, we mapped a trajectory within each person in terms of both AB and LS over the time spanning the pretest through the 3-month follow-up assessment. This provided us with our fundamental unit of analysis: a three-dimensional (3D; LS×AB×time) trajectory for each participant.

We then, in our second step, identified common trajectories in our samples through K-means cluster analysis [42] so that we might uncover the patterns of response change that naturally occur across this time span. Common trajectory shapes were identified in this way and individuals were grouped in these clusters based on their similarity with each of these shapes.

K-means cluster analysis [42] is an atheoretical approach to grouping people into K number of groups based purely on their similarity with one another. For this analysis, we performed a progressive series of K-means clustering, starting with K=1 and ending with K=10 and compared them to learn which value of K best represented our data (see the Multimedia Appendix 2 for a full description). The algorithm behind it employs prototypes to define each group, typically with the following computational steps (note also the differential usage of capitalization regarding the letter *K* to distinguish between the total number of clusters and the identification of a particular cluster).

1. Place K number of prototypes randomly within the reasonable observation window
2. Assign each subject to the prototype (k) it is closest to
3. For each k, reoptimize the prototype by defining it as the center (ie, mean) of all subjects that are assigned to it
4. Repeat steps 2 and 3 until none of the prototypes can be reasonably optimized further

XSL•FO
RenderX

Figure 1. Distance function used in the clustering of participant trends. Delta symbols are used to indicate differences between prototypes and individual trends.

$$distance = \int_{pre\ test}^{3mo\ follow\ up} \sqrt{\Delta f'(x)_{AB}{}^2 + \Delta f'(x)_{LS}{}^2}$$

Once each person was described by their own polynomial trajectory in both LS and AB over time (centered at posttest; the central-most observation in both samples), we executed a series of possible K-means clustering on these trajectories using the above steps. The major deviation in our approach from the typical K-means cluster analysis was that we employed a calculus-based distance function to quantify similarities in curve shape between trajectories, rather than the squared Euclidean distance between points, when assigning participants to prototypes and when reoptimizing prototypes to better represent the participants (this distance formula represents the limited integral of a Pythagorean combination of the difference between derivative functions within each predictor dimension; see Figure 1).

This deviation represents a modification of the K-means approach to match our research objectives: We are interested in how well-being changes over time, rather than well-being itself, and the purpose of our modification is to isolate the manner with which well-being changes. *Change*, in a mathematical sense, can be found by calculating the derivative of a function; the result is a function that plots the slope of the original trajectory, in this case, across time. By clustering along the total difference between derivatives (as we have done here), we have modified the standard K-means cluster analysis to cluster well-being trajectories based only on similarities in curve shape and ignoring original function intercepts.

With each participant clustered based on curve shape, we closed our analysis with an interpretation of these newly identified longitudinal outcomes (ie, clusters) and an analysis of predictors of group membership. Observing that our clusters inadvertently seemed to differentiate based on raw self-reported levels of well-being at baseline, we employed exploratory analyses of variance (ANOVAs) to describe group-wise summaries of responses across the observation period. We then used correlation and logistic regression to explore how depression, frequency of use, and adherence to instructions might relate to the distinct longitudinal outcomes and predict group membership.

Results

Common Trends

We began our analytic process by fitting a within-person polynomial regression model to both LS and AB scores across the available time span. Each of these bivariate models (2 for each participant) was raised to the polynomial degree of 2 less than the number of observations to produce precise, but unsaturated, representations of change in LS or AB over time. Our combination of models resulted in a 3D (ie, LS×AB×time) trajectory representation for each individual participant (Step 1). Trajectories were then grouped based on curve shape by performing the K-means cluster analysis [42] described above that employs a customized distance function to account for functional derivative differences rather than the Euclidean distance between points (Step 2). The final clusters were then interpreted through factorial ANOVA before being classified as distinct longitudinal outcomes and compared with one another in terms of self-reported depression symptoms and use statistics (Step 3).

The same approach was applied to both samples for confirmation and comparison purposes and all evaluations were initially conducted within the sample. Furthermore, as we discuss below and as indicated in Figures 2 and 3, similar trajectories were identified across samples in terms of both shape and position. In addition, taking into consideration the methodological and demographic similarities between the 2 samples, we decided to collapse our analysis of these variables across samples for concise presentation here except where stated.

XSL•FO
RenderX

Figure 2. Survival curves demonstrating the portion of participants retained over time.

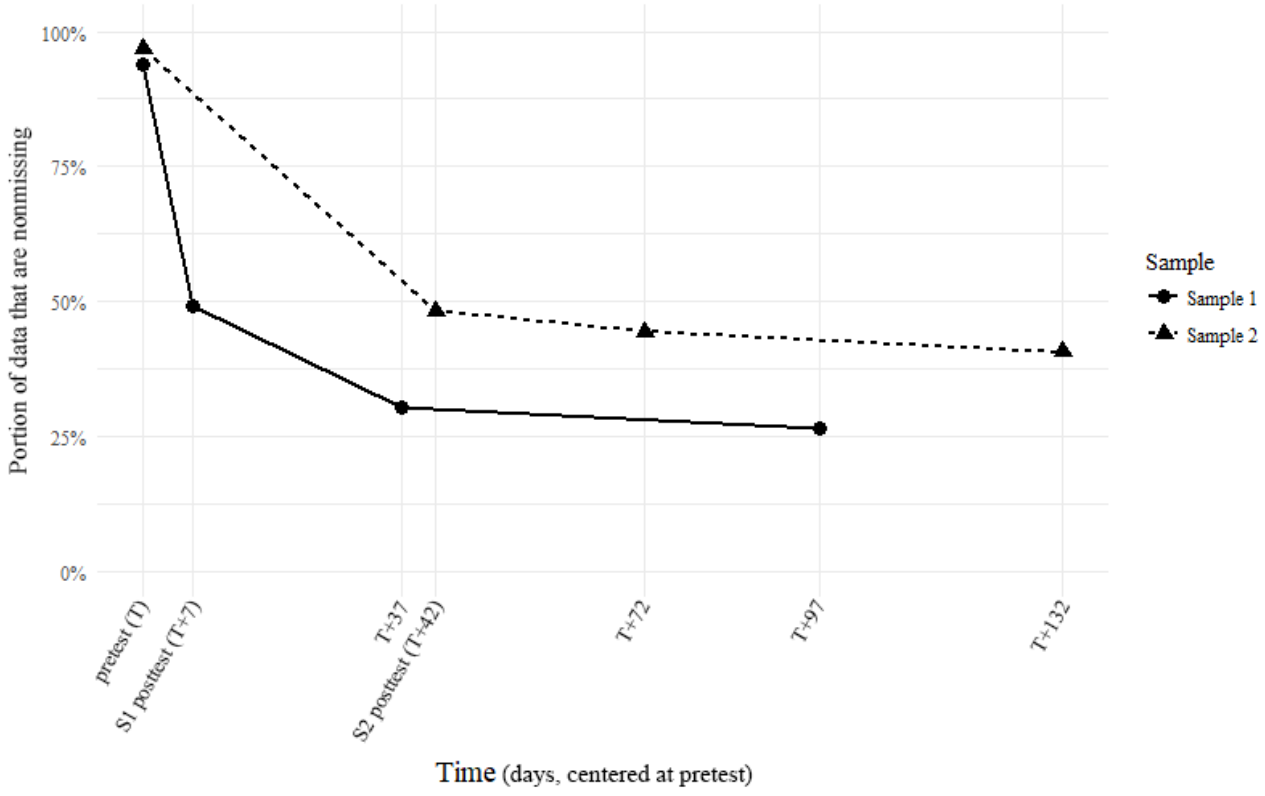
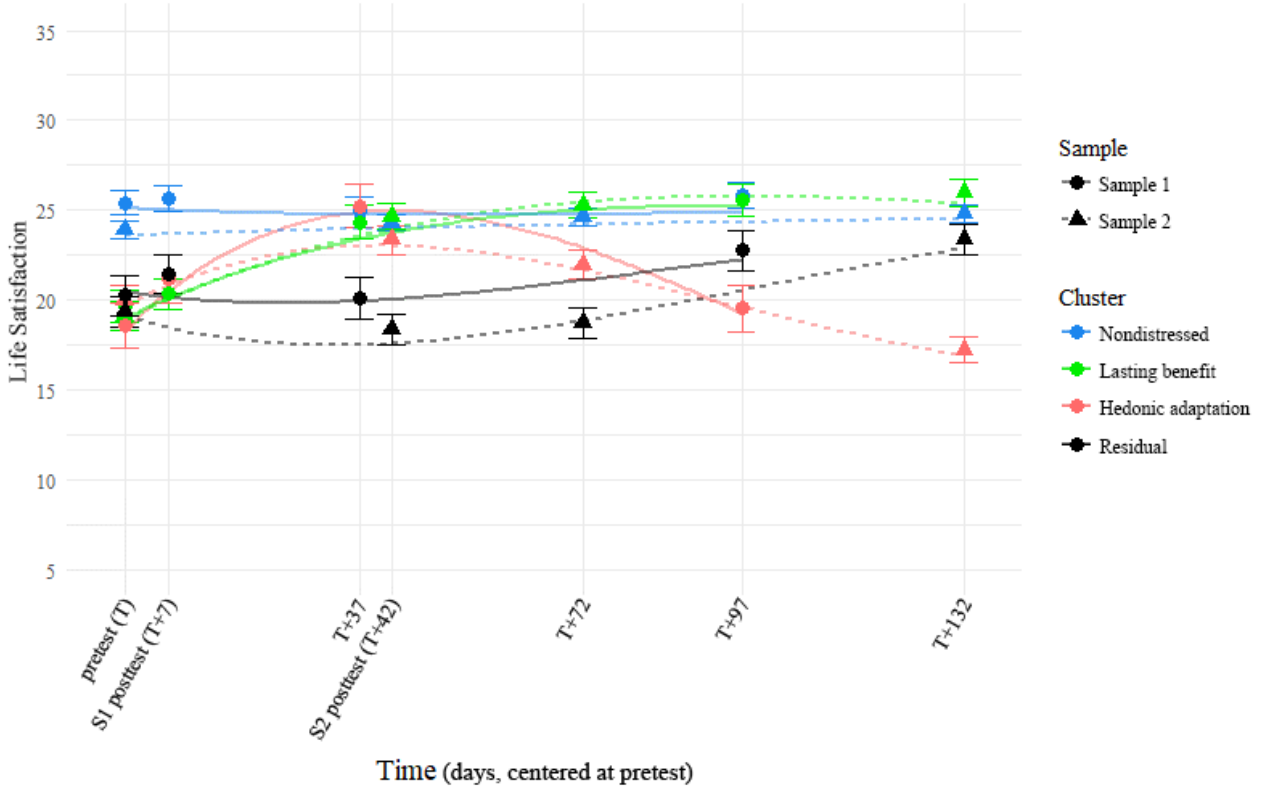


Figure 3. Life satisfaction trajectories over time by sample and cluster. The points in the foreground represent observed group-wise means, whereas the faded lines in the background represent the prototype trajectories that each cluster is based on. Error bars represent standard error of the mean.



Attrition

The primary concern in terms of attrition or missing data was the availability of data for our first step, the within-person modeling of happiness trajectories over time. Trajectories were mapped for all participants with at least 4 (from a possible total of 6) data points in both AB and LS between the pretest and the 1 year follow-up assessment. Although this approach produced

XSL•FO
RenderX

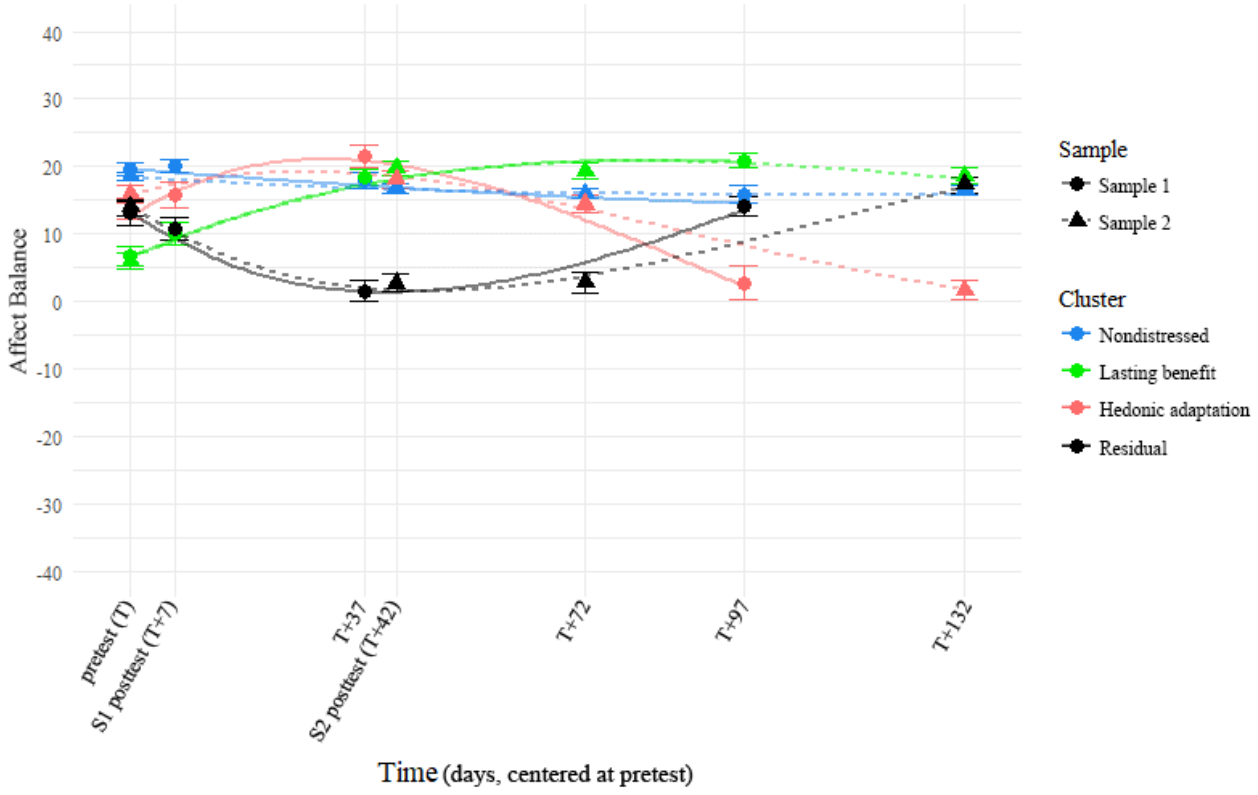
reasonable models of space between the pretest and the 3-month follow-up, it failed to reliably model the span of the following 9 months, wherein LS and AB were only assessed twice: 6 and 12 months postintervention. Despite many varied efforts, we were not able to arrive at any type of reasonable model for our data that included these time points.

These criteria reduced our sample sizes to 276 participants (30.3% of the original sample) in Sample 1 (where participants received only a 1-week intervention) and 579 participants (42.4%) in Sample 2, where the intervention was considerably more substantial. Proportionately, more participants qualified for analysis in Sample 2 than observed in Sample 1, χ^2_3=34.603, *P*<.001, potentially because of the difference in content between those 2 samples. Survival curves for the 2 samples are provided in Figure 2. Importantly, most of the attrition occurred between the first 2 assessments in either of the samples and this left open the possibility that the persons who dropped out did not complete, or even begin, their assigned intervention. Excluded participants reported significantly lower levels of LS (mean 20.49 [SD 8.15], n=1375), Welch's t_{1797}=-3.91, *P*<.001 and AB (mean 11.72 [SD 13.43], n=1382), Welch's t_{1913}=-5.74, *P*<.001, than the included participants (mean 21.89 and 14.92, respectively [SD 8.23 and 12.45, respectively], n=855) at pretest. They also reported a significantly higher prevalence of depression symptoms at this same time point (mean 18.22 [SD 12.97], n=1396), Welch's t_{1920}=5.24, *P*<.001, when compared with the participants who were retained in our analysis (mean 15.42 [SD 11.94]).

Trajectories in Life Satisfaction and Affect Balance Across Time

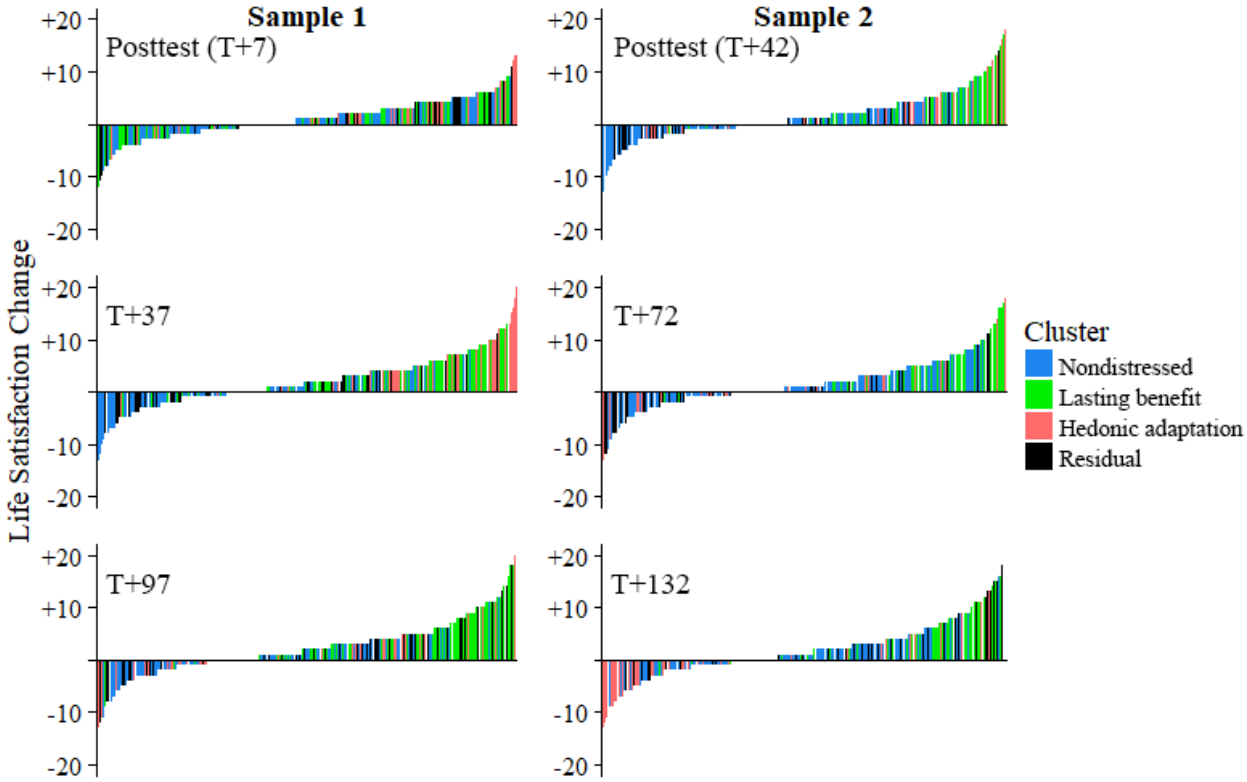
Using Cattell's [43] Scree approach and a descriptive comparison of the 4-factor solution to a 3- and 5-factor solution (see also Multimedia Appendix 2), we identified the 4 trajectory clusters described in this paper as being the most reasonable interpretations of trajectory shapes in the sample. These clusters are plotted in Figures 2 to 4 and were named *nondistressed* (42.8% of Sample 1 and 52.8% of Sample 2), *lasting benefit* (26.8% of Sample 1 and 18.0% of Sample 2), *hedonic adaptation* (13.4% of Sample 1 and 14.2% of Sample 2), and *residual* (17.0% of Sample 1 and 15.0% of Sample 2) based on the findings that we describe below. The reader should note that the observable similarity in common trajectory shapes across samples was not because of shared prototypes or algorithm seeds: separate and complete clustering were carried out within each of the 2 samples before displaying the results from both samples on the same plots here. Many of our subsequent comparisons were also carried out in this same way, with us occasionally reporting summary statistics across samples in the interest of brevity.

Figure 4. Affect balance trajectories over time by sample and cluster. The points in the foreground represent observed group-wise means, whereas the faded lines in the background represent the prototype trajectories that each cluster is based on. Error bars represent standard error of the mean.



XSL•FO
RenderX

Figure 5. Participant-level deviation in life satisfaction from baseline by sample, timepoint, and cluster. Each bar represents one participant (arranged by value and given pairwise deletion between plots). Raw within-person differences (from pretest) are represented on the y-axis. Similar information is represented in Figure 3, though it is presented here for visual confirmation of our cluster definitions.



It is worth noting that in addition to being atheoretical, our variation on this approach ignored function intercepts (and thereby the actual position of the original data points) in our attempt to classify trajectories purely in terms of their shape. Regardless, Figures 2 and 3 show that this clustering also seems to capture distinct group-wise difference in reporting at the various time points: Once an artificial intercept is applied to each prototype, it seems to lay directly on top of the observations from the cluster that the prototype represents. Figure 5 also shows that the group-wise summary statistics displayed in Figure 3 correspond well to the distributions of responses at each time point and further confirms the precision with which we captured distinct participant trajectories.

We first examined our forecast of group-wise differences in reporting using a separate 4×4 mixed factorial ANOVA for each of the 2 outcomes, LS and AB, in each of the 2 samples. In both samples, a significant timepoint (within-participant; pretest, posttest, 1-month follow-up, and 3-month follow-up) × cluster (between-participant; nondistressed, lasting benefit, hedonic adaptation, and residual) interaction effect qualified significant main effects of both timepoint and cluster (in Sample 1: significant 2-way interactions predicted both LS, $F_{9,645}$=19.67, P<.001, and AB, $F_{9,648}$=47.31, P<.001; in Sample 2: significant 2-way interactions predicted both LS, $F_{9,1611}$=53.412, P<.001, and AB, $F_{9,1611}$=95.697, P<.001; more information is available in Supplemental Tables 1 and 2). These significant interaction effects are direct evidence of mean-level differences between groups, so we examined them further with posthoc comparisons to learn where those difference occurred.

A full posthoc analysis of all possible pairwise comparisons within samples (with Bonferroni corrections) is presented in Supplemental Tables 1 and 2. In summary, we found strong statistical evidence for exactly the patterns we would expect by observing Figures 2,3, and 4. First, a chronically high-reporting group (later identified as the nondistressed cluster; Sample 1 n=118, Sample 2 n=306) emerged across samples. As can be seen in supplemental tables 1 and 2, membership in this cluster significantly predicted higher levels of both LS and AB than the other 3 groups at the pretest assessment (T=0). Second, we observed a group of individuals in both samples who seemed to lastingly benefit from a given positive intervention (the lasting benefit cluster; Sample 1 n=74, Sample 2 n=104). Although these persons began the intervention with levels of LS and AB at or below the 3 other clusters, they showed a marked improvement in these regard when assessed a little over a month later (T+37 in Sample 1 and T+42 in Sample 2), regardless of intervention length. Furthermore, the K-means clustering approach also identified a group in each sample that showed a similar benefit around the same time point, but then displayed a hedonic adaptation curve afterward (the hedonic adaptation cluster; Sample 1 n=37 and Sample 2 n=82). Significant differences in the reporting of LS and AB between this cluster and the nondistressed and lasting benefit clusters can be observed around the 3-month follow-up in both samples. Lastly, a group of persons emerged with each clustering which we found difficult to explain in this context (the residual cluster; Sample 1 n=47 and Sample 2 n=87). Persons in this group did not seem to report an immediate benefit because of the intervention and displayed significantly less LS and AB than the other 3 groups at 3 time points: 37 (T+37; Sample 1), 42

XSL•FO
RenderX

(T+42; Sample 2), and 72 (T+72; Sample 2) days after the start of the intervention. Across separate clustering in 2 different samples, our approach arrived at what seem to be the same 4 distinct clusters of trajectories, regardless of the major differences in the interventions and timeframes between the 2 samples.

A posthoc comparison of reported LS and AB values between samples is made difficult by the fact that when the time frame is linked at the pretest (ie, when we consider the pretest assessment as our starting point), the later assessment timepoints are no longer equitable across samples. Our argument for consistency across samples in the clusters formed, however, requires some means of evaluating the similarity in cluster solutions between samples. To this end, we chose to compare reported LS and AB values between samples within each of the following 3 time point groups: early assessments (T and T+7), middle assessments (T+37 and T+42), and late assessments (T+72, T+97, and T+132; see also Supplemental Tables 1 and 2 for group-wise means and SDs by timepoint and sample). When accounting for cluster-wise and sample-wise main effects with Type 3 sums of squares, we did not observe a significant interaction effect between cluster and sample during the early assessments in terms of either LS, $F_{3, 971}$=0.60, P=.62, or AB, $F_{3,971}$=1.76, P=.15. This lack of a significant interaction effect was again observed within the middle assessments in terms of both LS, $F_{3,752}$=0.51, P=.67, and AB, $F_{3,752}$=1.12, P=.34, and again within the late assessments of LS, $F_{3,1293}$=0.44, P=.73. A significant interaction effect between cluster and sample did emerge during the late assessments of AB, $F_{3,1293}$=3.05, P=.028, though this effect accounted for less than 1% of the total variance in the outcome, η^2=.006. We contend that the general lack of an observable statistical effect is not equivalent to observance of nondifference; however, we find the qualitative similarities observable in Figures 1 and 2 to be supported by this body of null findings.

Table 1. General binomial (logistic) models predicting membership in the lasting benefit cluster over the hedonic adaptation cluster from use statistics.

Factor	Model 1	Model 2	Model 3
Fixed effects	Intercept	Intercept	Intercept
	time[a]	time	time
	freq[b]	freq	freq
	adh[c]	adh	adh
	freq×[d]adh	freq×adh	freq×adh
	—[e]	freq×time	freq×time
	—	—	adh×time
	—	—	freq×adh×time
Model *df*	5	6	8
Log likelihood	−877.23	−873.28	−872.92
R-squared[f]	[.570,.641]	[.572,.644]	[.573,.645]
Akaike information criterion	1764.5	1758.6	1761.8
Residual deviance	1754.5	1746.6	1745.8
Residual *df*	1295	1294	1292
Comparison with previous model			
χ^2	—	7.9	0.73
df	—	1	2
P value	—	0.005	0.7

[a]Assessment time point, in days (possible values before mean centering: 7, 14, 21, 28, 35, 37, 42, 72, 97, and 132).

[b]Frequency of use (in number of days per week; scaled by total intervention length).

[c]Adherence to the specific instructions of an exercise within the past week (coded as 1=true, 0=false).

[d]Indicates an interaction effect.

[e]Not applicable.

[f]The limits of the R^2 statistics presented here are Cox & Snell's pseudo- R^2 and Nagelkerke's pseudo- R^2, respectively.

XSL•FO
RenderX

Table 2. Optimal binomial model to describe the relationship between intervention use and the achievement of a lasting intervention benefit.

Fixed effects	Estimate (*b*)	Standard error	Standardized (β)	*z* score	*P* value
Intercept	0.47	0.11	.00	4.45	<.001
freq[a]	-0.20	0.10	-.35	-1.97	<.05
adh[b]	0.11	0.24	.09	0.45	.66
time[c]	-0.00	0.00	-.05	-0.37	.72
freq×adh	0.65	0.23	.54	2.88	.004
freq×time	0.01	0.00	.36	2.75	.006

[a]Frequency of use (scaled by intervention length).

[b]Adherence to the specific instructions of an exercise within the past week (coded as 1=true, 0=false).

[c]Assessment time point, in days (range of values before mean centering: 7, 14, 21, 28, 35, 37, 42, 72, 97, and 132).

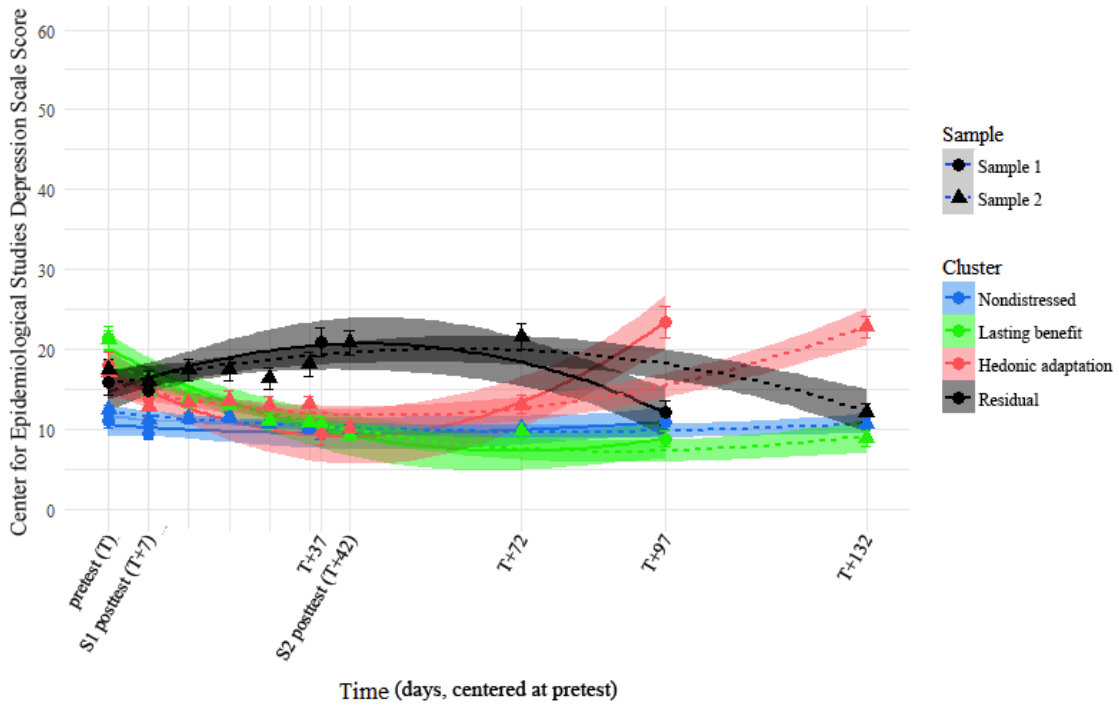
Relationships Between Identified Longitudinal Trajectories and Other Variables

Next, we approach the question of what might have caused some people to exhibit one pattern over another (ie, the third step of our analytic plan). In an attempt to address this question, we employed a handful of logistic regression models predicting cluster membership from metrics of depression symptomology (specifically, CES-D scores) and metrics relating to the use and continued use of the assigned activity or activities.

Depression

The presence of depression symptoms was assessed with the CES-D [36] at all the same timepoints where LS and AB were measured, as well as weekly during the intervention period in Sample 2. This allowed us to explore complex multidimensional multilevel models in our assessment of how this metric differentiated the 4 clusters, though the findings are relatively straightforward and can be best summarized simply by the correlations between variables: Depression was strongly negatively correlated with both LS, r (n=5076)=−.70, P<.001, 95% CI −0.71 to −0.70, and AB, r (n=5084)=−.75, P<.001, 95% CI −0.76 to −0.75, and these 2 outcomes were strongly correlated with one another, r (n=5076)=.62, P<.001, 95% CI 0.60 to 0.62. The parallels between Figures 2 and 3 are again replicated when plotting depression over this same time frame: The changes that we observe over time are not only limited to LS and AB, Figure 6 demonstrates that they also encompass large-scale changes in the presence of depression symptoms over time.

Figure 6. Trends in self-reported depression symptoms over time by sample and cluster. Group-wise means and standard errors are represented by points and error bars; trajectory curves are formed using a Loess smoothing function with a span width of 2 days. The standard error of the smoothing function is represented by shaded regions.



XSL•FO
RenderX

This finding emerged, in part, because of 2 factors. First, nearly half of the participants (1051/2241, 46.9% responses before deletion) began the intervention with self-reported depression symptomology that met or exceeded the common preclinical criterion for a high risk of depression (a score of 16 or above on the CES-D [36]). Second, the self-reported changes in LS and AB, similar to changes in depression symptoms, were simply more profound in distressed participants than the changes in LS and AB that occurred within seemingly nondistressed participants. With these findings, we solidified our categorization of the most popular trajectory across samples as a trajectory specific to distinctly nondistressed individuals; the remaining clusters, and the focus of our investigation are the distressed individuals among which we can more easily identify longitudinal change.

Frequency of Use and Adherence to Instructions

Use and adherence measures were assessed for each exercise that the participant was administered, which remained at 1 exercise for participants in Sample 1 and varied between 0 and 6 exercises in Sample 2. This became a problem for the assessment of how frequently the participant used that exercise over the past week, so these use metrics were grouped based on intervention length and z-standardized within-group. Given our findings from the depression analysis, we employed frequency of use and adherence to instructions as possible predictors in distinguishing between participants who showed a lasting benefit because of the intervention (n=146,1206 assessments) as opposed to a hedonic adaptation curve during the assessment period (n=92,886 assessments).

In searching for an optimally fitting model to describe our data, we began with a maximal multilevel binomial model predicting a lasting benefit over hedonic adaptation and sequentially removed predictors from this model based on overall model fit. More specifically, the original maximal model predicted the binomial outcome (0=hedonic adaptation, 1=lasting benefit) from the interaction of time (days; mean centered around 28.36)×use (standardized as described above)×adherence (binomial: 0=did not adhere to the specific exercise instructions, 1=did adhere) and all lower-order interaction and main effects that comprise this 3-way interaction, with slopes and intercepts varied by participant, condition, and sample. All random effects were removed based on sequential comparisons of models, as was the 3-way interaction effect and the 2-way interaction effect between adherence and time. The final steps of this process are summarized in Table 1.

Table 1 displays summary fit statistics and comparisons between some of the latter models we considered. The outcome of each of the above functions is membership in the lasting benefit cluster (coded as *1*; functional n=1206) over membership in the hedonic adaptation cluster (coded as *0*; functional n=886). Participants were drawn from both samples and all data points observed between the pretest and the 3-months follow-up assessment are included. All predictors represented here, including time, have been grand mean centered. Model 2 is observed to be an optimal description for these data.

A closer examination of the final (optimal) model is provided in Table 2. Looking at the parameter estimates for our optimal model, we can see a number of key relationships between frequency of use, adherence to instructions, and the attainment of a lasting benefit rather than hedonic adaptation. For those who adhered to the specific intervention instructions, frequency of use considerably predicted a lasting benefit across all time points, with an especially large effect at later assessments. For those who reported that they were not adhering to the instructions of their intervention, frequency of use made relatively no impact on their possibility of membership in the lasting benefit cluster; these participants generally exhibited about a 50% probability of belonging to either group.

Table 2 presents the parameter estimates of the best-fitting model of the relationship between use statistics and membership in the lasting benefit (rather than hedonic adaptation) cluster (ie, the model labelled Model 2 in Table 1). Functional N=2092. McFadden's R^2=.387. Deviance residuals: minimum=−1.78, first quartile=−1.28, median=0.95, third quartile=1.03, maximum=1.28.

Discussion

Principal Findings

This analysis examined how well-being changes over time for people enrolled in an OPPI. Not surprisingly, in light of previous research showing that high baseline well-being predicts smaller OPPI effects [44], we found that the individual's well-being at baseline had a major impact on these trajectories in our samples. Persons who were relatively well-off at the start of the intervention reported much smaller changes over time than the distinctly distressed persons who comprised roughly half of our samples. In terms of each of our specific research questions (presented in the introduction of the present paper), respectively:

1. For the distressed portion of participants, our cluster analysis of trajectories revealed that a lasting benefit following the OPPI might be just as likely as a temporary benefit (ie, adaptation); both experiences were commonly reported by participants.
2. The distressed participants who continued to use the OPPI were much more likely to see a lasting benefit rather than adaptation, especially at later time points, but only if they also reported adhering to the specific instructions of whatever OPPI they were assigned to.
3. A substantial portion of the distressed participants also exhibited trajectories that defied classification as either a lasting benefit or adaptation, though these individuals typically did not report any immediate benefit of the OPPI between the pretest and posttest. For those who were not apparently distressed at baseline, the changes to well-being over time were much subtler than the changes exhibited by distressed participants and thus more difficult to identify given the current approach.

Participants were randomly assigned between a total of 18 different conditions between the 2 samples, with 1 group representing a waitlist control condition, another group receiving a Web-based placebo intervention, and the remaining 16 experimental conditions receiving an OPPI (see Multimedia Appendix 1 for specific exercise instructions). Despite these

XSL•FO
RenderX

OPPIs varying in both dosage and content, we were unable to observe any differences in participant outcomes because of condition assignment. A number of factors could explain this lack of a finding, but we contend that it is likely more because of the limitations of our design and samples rather than a general inefficacy of OPPIs. These limitations are explored further below, followed by recommendations for the design of OPPIs and OPPI studies.

Limitations

This analysis has a number of limitations that stem from the manner with which data were collected. Our data drew from 2 samples of OPPIs, both collected from a website that appealed to those seeking to increase their happiness. Participants joined with little incentive for participation and many were recruited using an advertisement that appeared at the end of a happiness self-help book. Thus, it is hard to isolate the impact of any one specific exercise. OPPIs tend to attract motivated happiness seekers who might use additional Web-based or offline resources to boost their happiness. Our findings can also only generalize to these types of people; arguments are made here and elsewhere [19,44] as to the ways that Web-based happiness seekers differ from other members of the public. It is also worth noting, however, that this context mirrors many ways in which such interventions are deployed in real-world settings. Few incentives exist to encourage individuals to engage in such interventions; their own self-interest and happiness is just one example of something people might be interested in changing but other examples could be mental distress, physical activity, and weight.

Although our findings might suggest that users with high initial well-being derive less benefit from OPPIs, it is important to note that this pattern of findings may or may not generalize other outcomes not measured in this study. For example, high well-being users may gain resilience against future stress, or greater self-efficacy about dealing with future stress, neither of which is measured in the 2 study samples. In addition, our analysis is limited to the span of the first few months after opting into an OPPI and is based on no more than a handful of assessments per individual. Internal data collected by the third author (A. Parks) reveals that higher well-being users on an interactive well-being platform must apply more effort to achieve the same effect as their low well-being counterparts, and as a result, often take longer to reach the same level of well-being improvement. It is possible, therefore, that what appears to be a smaller effect is just an effect with longer latency or one that could be achieved if high well-being users were offered more activities to use. Finally, both samples offered participants a constrained selection of activities, which is not consistent with real-world settings, where users have many available options and can choose freely between them [45].

Recommendations

A publication by the third author [46] outlines a number of key ways in which OPPIs and OPPI studies can be improved to maximize their efficacy and meet the ever-evolving best practices in eHealth intervention research. These findings echo support for those recommendations by demonstrating the importance of continuing to be engaged with and adherent to an OPPI long after the OPPI is learnt. Thus, we review here those recommendations that are most pertinent to user engagement and adherence.

Generally, eHealth and mHealth interventions suffer from a great deal of attrition and interventionists can potentially combat this by increasing user engagement and allowing free choice in how one self-improves [46]. The use of Web-based or mobile platforms also allows for the customization of intervention content to each user, which increases compatibility and can subsequently lead to greater intervention adherence [47]. Furthermore, the benefits of practicing an eHealth or mHealth intervention can be made apparent to the user by using technology to automatically provide quantitative feedback on a user's progress over time and how they compare to other users. Although freedom of choice and the provision of feedback can generate problems for intervention researchers, these are the types of things that we would expect to be most beneficial to users in terms of encouraging engagement, retention, and adherence to the Web-based intervention.

There is still a great deal of work to be done in OPPI research, and we continue to echo the previous recommendations of Parks [46] as we offer guidance for this research based on these findings and their limitations. This analysis goes beyond the traditional approach of reporting average group differences to explore how participant outcomes vary across time regardless of any preconceived notion of how participants should be grouped, and we have hopefully been successful in demonstrating the value of such an approach in complementing other OPPI research. Again, attrition is a severe problem in eHealth research [48], as it is in this report, and we recommend that future investigators commit resources to aggressive email and phone follow-up with a subset of persons who drop out of a study to learn if those persons are continuing to use the intervention and utilize that information in interpreting and handling missing data. Although we acknowledge that it is important to observe dropout rates in naturalistic settings, rather than artificially driving up retention with monetary incentives, we nevertheless acknowledge the statistical difficulties that arise when large numbers of participants are missing and encourage researchers to think of creative solutions to reduce dropout and to learn more about those who do drop out.

Conclusions

Persons suffering from moderate to severe depression symptoms are naturally attracted by advertisements for Web-based happiness-promoting intervention studies. This study provides additional evidence that OPPIs can support lasting benefits among this type of user, especially those who show regular usage. More research is needed using a greater variety of outcome measures to determine whether the impact is indeed smaller on high well-being users, or if the benefits conferred on high well-being users simply are not detectable using the currently reported outcomes. Partially explaining this recovery, those who continued to practice the activities that they received in these 2 studies continued to see gains from them 1 and 3 months after the intervention ended. The major contribution of this paper, however, is the demonstration of a novel approach within Web-based interventions to investigating long-term benefits leveraging person-centric analytic methods. Many

XSL•FO
RenderX

people who might be interested in such resources would likely be interested in knowing whether such an intervention would likely benefit them. A better understanding of who benefits from Web-based interventions might also improve their use in stepped care programs where low-intensity Web-based interventions can be offered to those individuals in efforts to save cost and improve population health. Considering these findings, we advocate for the continued study of OPPIs for depressive symptoms—perhaps as an adjunct to standard treatments—and to facilitate optimal long-term outcomes, we encourage the design of OPPIs that maximize continued engagement and adherence after the intervention is over.

Acknowledgments

The authors extend their sincere gratitude to the Positive Psychology Center at the University of Pennsylvania, where data analyzed in this paper were gathered as well as the countless experts who have helped in the development of this study.

Conflicts of Interest

The authors declare no conflicts of interest.

Multimedia Appendix 1

Complete listing of exercise descriptions and schedule.

[PDF File (Adobe PDF File), 62KB - jmir_v21i6e13253_app1.pdf]

Multimedia Appendix 2

Further explanation of our derivation of a four-factor solution.

[PDF File (Adobe PDF File), 176KB - jmir_v21i6e13253_app2.pdf]

References

1. Leykin Y, Aguilera A, Torres LD, Pérez-Stable EJ, Muñoz RF. Interpreting the outcomes of automated internet-based randomized trials: example of an International Smoking Cessation Study. J Med Internet Res 2012 Feb;14(1):e5 [FREE Full text] [doi: 10.2196/jmir.1829] [Medline: 22314016]
2. Gow RW, Trace SE, Mazzeo SE. Preventing weight gain in first year college students: an online intervention to prevent the “freshman fifteen”. Eat Behav 2010 Jan;11(1):33-39 [FREE Full text] [doi: 10.1016/j.eatbeh.2009.08.005] [Medline: 19962118]
3. Hobfoll SE, Blais RK, Stevens NR, Walt L, Gengler R. Vets prevail online intervention reduces PTSD and depression in veterans with mild-to-moderate symptoms. J Consult Clin Psychol 2016 Jan;84(1):31-42. [doi: 10.1037/ccp0000041] [Medline: 26322788]
4. Bolier L, Abello KM. Online positive psychological interventions: state of the art and future directions. In: Parks AC, Schueller SM, editors. The Wiley Blackwell Handbook of Positive Psychological Interventions. Oxford, England: Wiley Blackwell; 2014:286-309.
5. Muñoz AF, Chavira DA, Himle JA, Koerner K, Muroff J, Reynolds J, et al. Digital apothecaries: a vision for making health care interventions accessible worldwide. mHealth 2018 Jun;4:18. [doi: 10.21037/mhealth.2018.05.04]
6. Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. Proc Natl Acad Sci U S A 2018 Dec 3;115(27):E6106-E6115. [doi: 10.1073/pnas.1711978115] [Medline: 29915059]
7. Goldstein ED. Sacred moments: implications on well-being and stress. J Clin Psychol 2007 Oct;63(10):1001-1019. [doi: 10.1002/jclp.20402] [Medline: 17828765]
8. Mitchell J, Stanimirovic R, Klein B, Vella-Brodrick D. A randomised controlled trial of a self-guided internet intervention promoting well-being. Comput Hum Behav 2009 May;25(3):749-760. [doi: 10.1016/j.chb.2009.02.003]
9. Seligman ME, Steen TA, Park N, Peterson C. Positive psychology progress: empirical validation of interventions. Am Psychol 2005 Aug;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
10. Shapira LB, Mongrain M. The benefits of self-compassion and optimism exercises for individuals vulnerable to depression. J Posit Psychol 2010 Sep;5(5):377-389. [doi: 10.1080/17439760.2010.516763]
11. Parks-Sheiner AC. Academia.edu.: ProQuest Information & Learning; 2009. Positive psychotherapy: Building a model of empirically supported self-help URL: http://www.academia.edu/download/31388758/parksdissertationFINALacknow.pdf [accessed 2019-06-03] [WebCite Cache ID 78rIEugjr]
12. Bolier L. [Positive psychology online]. In: Bohlmeijer E, Bolier L, Westerhof G, Walburg JA, editors. [Handbook Positive Psychology: Theory, Research, Applications]. Amsterdam: Uitgeverij Boom; 2013:355-371.
13. Gander F, Proyer RT, Ruch W, Wyss T. Strength-based positive interventions: further evidence for their potential in enhancing well-being and alleviating depression. J Happiness Stud 2012;14:1241-1259. [doi: 10.1007/s10902-012-9380-0]

XSL•FO
RenderX

14. Lyubomirsky S. Why are some people happier than others? The role of cognitive and motivational processes in well-being. Am Psychol 2001;56(3):239-249. [doi: 10.1037/0003-066X.56.3.239]
15. Bolier L, Haverman M, Westerhof GJ, Riper H, Smit F, Bohlmeijer E. Positive psychology interventions: a meta-analysis of randomized controlled studies. BMC Public Health 2013 Feb 8;13:119 [FREE Full text] [doi: 10.1186/1471-2458-13-119] [Medline: 23390882]
16. Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. J Clin Psychol 2009 May;65(5):467-487. [doi: 10.1002/jclp.20593] [Medline: 19301241]
17. Layous K, Nelson SK, Lyubomirsky S. What is the optimal way to deliver a positive activity intervention? The case of writing about one's best possible selves. J Happiness Stud 2012 May 6;14(2):635-654. [doi: 10.1007/s10902-012-9346-2]
18. Lyubomirsky S, Dickerhoof R, Boehm JK, Sheldon KM. Becoming happier takes both a will and a proper way: an experimental longitudinal intervention to boost well-being. Emotion 2011 Apr;11(2):391-402 [FREE Full text] [doi: 10.1037/a0022575] [Medline: 21500907]
19. Parks AC, Della Porta MD, Pierce RS, Zilca R, Lyubomirsky S. Pursuing happiness in everyday life: the characteristics and behaviors of online happiness seekers. Emotion 2012 Dec;12(6):1222-1234. [doi: 10.1037/a0028587] [Medline: 22642345]
20. Haeck C, Schueller SM, Parks AC. Impact of psychoeducational content delivered online to a positive psychology aware community. J Posit Psychol 2016;11(3):270-275 [FREE Full text] [doi: 10.1080/17439760.2015.1048819] [Medline: 27499798]
21. Kaczmarek LD, Kashdan TB, Kleiman EM, Baczkowski B, Enko J, Siebers A, et al. Who self-initiates gratitude interventions in daily life? An examination of intentions, curiosity, depressive symptoms, and life satisfaction. Pers Individ Dif 2013 Oct;55(7):805-810. [doi: 10.1016/j.paid.2013.06.013]
22. Lykken D, Tellegen A. Happiness is a stochastic phenomenon. Psychol Sci 1996;7(3):186-189. [doi: 10.1111/j.1467-9280.1996.tb00355.x]
23. Brickman P, Campbell DT. Hedonic relativism and planning the good society. In: Appley MH, editor. Adaptation-Level Theory: A Symposium. New York: Academic Press; 1972:287-305.
24. Frederick S, Loewenstein G. Hedonic adaptation. In: Kahneman D, Diener E, Schwartz N, editors. Well-Being: Foundations Of Hedonic Psychology. New York: Russell Sage Foundation; 1999:302-329.
25. Lybomirsky S. Hedonic adaptation to positive and negative experiences. In: Folkman S, editor. The Oxford Handbook Of Stress, Health, And Coping. New York: Oxford University Press; 2012:200-224.
26. Brickman P, Coates D, Janoff-Bulman R. Lottery winners and accident victims: is happiness relative? J Pers Soc Psychol 1978 Aug;36(8):917-927. [Medline: 690806]
27. Chadi A, Hetschko C. The magic of the new: how job changes affect job satisfaction. J Econ Manage Strat 2017 Jul 25;27(1):23-39. [doi: 10.1111/jems.12217]
28. Dyrdal GM, Lucas RE. Reaction and adaptation to the birth of a child: a couple-level analysis. Dev Psychol 2013 Apr;49(4):749-761. [doi: 10.1037/a0028335] [Medline: 22686177]
29. Kramrei E, Coit C, Martin S, Fogo W, Mahoney A. Post-divorce adjustment and social relationships: a meta-analytic review. J Divor Remarr 2007 Feb 23;46(3-4):145-166. [doi: 10.1300/J087v46n03_09]
30. Sheldon KM, Lyubomirsky S. The challenge of staying happier: testing the Hedonic Adaptation Prevention model. Pers Soc Psychol Bull 2012 May;38(5):670-680. [doi: 10.1177/0146167212436400] [Medline: 22361725]
31. Wilson TD, Gilbert DT. Explaining away: a model of affective adaptation. Perspect Psychol Sci 2008 Sep;3(5):370-386. [doi: 10.1111/j.1745-6924.2008.00085.x] [Medline: 26158955]
32. Lyubomirsky S, Sheldon KM, Schkade D. Pursuing happiness: the architecture of sustainable change. Rev Gen Psychol 2005;9(2):111-131. [doi: 10.1037/1089-2680.9.2.111]
33. Sheldon KM, Boehm JK, Lyubomirsky S. Variety is the spice of happiness: the Hedonic Adaptation Prevention (HAP) model. In: Boniwell I, David SA, Ayers AC, editors. The Oxford Handbook Of Happiness. Oxford: Oxford University Press; 2012:901-914.
34. Lyubomirsky S, Layous K. How do simple positive activities increase well-being? Curr Dir Psychol Sci 2013 Jan 22;22(1):57-62. [doi: 10.1177/0963721412469809]
35. Schueller SM, Parks AC. Disseminating self-help: positive psychology exercises in an online trial. J Med Internet Res 2012 Jun;14(3):e63 [FREE Full text] [doi: 10.2196/jmir.1850] [Medline: 22732765]
36. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 1;1(3):385-401. [doi: 10.1177/014662167700100306]
37. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
38. Pavot W, Diener E. The Satisfaction With Life Scale and the emerging construct of life satisfaction. J Posit Psychol 2008 Apr;3(2):137-152. [doi: 10.1080/17439760701756946]
39. Watson D, Clark LA, McIntyre CW, Hamaker S. Affect, personality, and social activity. J Pers Soc Psychol 1992 Dec;63(6):1011-1025. [Medline: 1460554]

XSL•FO
RenderX

40. Stagl JM, Antoni MH, Lechner SC, Bouchard LC, Blomberg BB, Glück S, et al. Randomized controlled trial of cognitive behavioral stress management in breast cancer: a brief report of effects on 5-year depressive symptoms. Health Psychol 2015 Feb;34(2):176-180. [doi: 10.1037/hea0000125] [Medline: 25068452]
41. Vilagut G, Forero CG, Barbaglia G, Alonso J. Screening for depression in the general population with the Center for Epidemiologic Studies Depression (CES-D): a systematic review with meta-analysis. PLoS One 2016;11(5):e0155431 [FREE Full text] [doi: 10.1371/journal.pone.0155431] [Medline: 27182821]
42. Hartigan JA, Wong MA. Algorithm AS 136: a k-means clustering algorithm. Appl Stat 1979;28(1):100-108. [doi: 10.2307/2346830]
43. Cattell RB. The scree test for the number of factors. Multivariate Behav Res 1966 Apr;1(2):245-276. [doi: 10.1207/s15327906mbr0102_10]
44. Sergeant S, Mongrain M. Distressed users report a better response to online positive psychology interventions than nondistressed users. Can Psychol 2015;56(3):322-331. [doi: 10.1037/cap0000034]
45. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the. J Med Internet Res 2016 Aug 31;18(8):e241 [FREE Full text] [doi: 10.2196/jmir.5725] [Medline: 27580524]
46. Parks AC. A case for the advancement of the design and study of online positive psychological interventions. J Posit Psychol 2014 Jul 11;9(6):502-508. [doi: 10.1080/17439760.2014.936969]
47. 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;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
48. 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]

Abbreviations

AB: affect balance
ANOVA: analysis of variance
CES-D: Center for Epidemiological Studies–depression scale
eHealth: electronic health
IRB: Institutional Review Board
LS: life satisfaction
mHealth: mobile health
OPPI: online positive psychological intervention
PPI: positive psychological intervention

Edited by G Eysenbach; submitted 28.12.18; peer-reviewed by J Marcotte, R Choubisa; comments to author 23.03.19; revised version received 19.04.19; accepted 14.05.19; published 08.06.19.

Please cite as:
Sanders CA, Schueller SM, Parks AC, Howell RT
Understanding Long-Term Trajectories in Web-Based Happiness Interventions: Secondary Analysis From Two Web-Based Randomized Trials
J Med Internet Res 2019;21(6):e13253
URL: http://www.jmir.org/2019/6/e13253/
doi:10.2196/13253
PMID:31199342

©Christopher A Sanders, Stephen M Schueller, Acacia C Parks, Ryan T Howell. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 08.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation of a Web-Based Work-Related Psychological Aftercare Program Into Clinical Routine: Results of a Longitudinal Observational Study

Rüdiger Zwerenz[1], Dr Biol Hom; Carlotta Baumgarten[1], MSc; Ingo Dahn[2], Dr Sc; Nicole Labitzke[3], PhD; Andreas Schwarting[4], Dr Med; Matthias Rudolph[5], Dr Med; Peter Ferdinand[2], PhD; Ute Dederichs-Masius[6], Dr Med; Manfred E Beutel[1], Dr Med

[1]Department for Psychosomatic Medicine and Psychotherapy, University Medical Center, Johannes Gutenberg-University, Mainz, Germany

[2]Knowledge Media Institute, University of Koblenz-Landau, Koblenz, Germany

[3]Center for Audiovisual Production, Johannes Gutenberg-University, Mainz, Germany

[4]Center of Rheumatology, University Medical Center, Johannes Gutenberg-University, Mainz, Germany

[5]Mittelrhein-Klinik, German Statutory Pension Insurance Rhineland-Palatinate, Boppard, Bad Salzig, Germany

[6]Drei-Burgen-Klinik, German Statutory Pension Insurance Rhineland-Palatinate, Bad Kreuznach, Germany

Corresponding Author:
Rüdiger Zwerenz, Dr Biol Hom
Department for Psychosomatic Medicine and Psychotherapy
University Medical Center
Johannes Gutenberg-University
Untere Zahlbacher Str 8
Mainz, 55131
Germany
Phone: 49 6131175981
Fax: 49 613117475981
Email: ruediger.zwerenz@unimedizin-mainz.de

Abstract

Background: As inpatient medical rehabilitation serves to promote work ability, vocational reintegration is a crucial outcome. However, previous Web-based trials on coping with work-related stress have been limited to Web-based recruitment of study participants.

Objective: The aim of our study was to evaluate the implementation of an empirically supported transdiagnostic psychodynamic Web-based aftercare program *GSA* (*Gesund und Stressfrei am Arbeitsplatz* [Healthy and stress-less at the workplace])-*Online plus* into the clinical routine of inpatient medical rehabilitation, to identify characteristics of patients who have received the recommendation for *GSA-Online plus*, and to determine helpfulness of the intervention and satisfaction of the participants as well as improvement in quality of life and mental health status of the regular users of *GSA-Online plus*.

Methods: GSA-Online plus was prescribed by physicians at termination of orthopedic psychosomatic inpatient rehabilitation. Participants' use of the program, work-related attitudes, distress, and quality of life were assessed on the Web.

Results: In 2 rehabilitation centers, 4.4% (112/2562) of rehabilitants got a recommendation for *GSA-Online plus* during inpatient rehabilitation. Compared with usual person aftercare, the Web-based aftercare program was rarely recommended by physicians. Recommendations were made more frequently in psychosomatic (69/1172, 5.9%) than orthopedic (43/1389, 3.1%) rehabilitation (χ^2_1=11.845, *P*=.001, *Cramér V*=−0.068) and to younger patients (*P*=.004, *d*=0.28) with longer inpatient treatment duration (*P*<.001, *r*=−0.12) and extended sick leaves before inpatient medical rehabilitation (*P*=.004; *Cramér V*=0.072). Following recommendation, 77% (86/112) of rehabilitants participated in Web-based aftercare. Completers (50/86, 58%) reported statistically significant improvements between discharge of inpatient treatment and the end of the aftercare program for subjective work ability (*P*=.02, *d*=0.41), perceived stress (*P*=.01, *d*=−0.38), functioning (*P*=.002, *d*=−0.60), and life satisfaction (*P*=.008, *d*=0.42).

Conclusions: Physicians' recommendations of Web-based aftercare are well accepted by patients who derive considerable benefits from participation. However, a low rate of prescription compared with other usual aftercare options points to barriers among physicians to prescribing Web-based aftercare.

XSL•FO
RenderX

(J Med Internet Res 2019;21(6):e12285) doi:10.2196/12285

KEYWORDS

aftercare; internet; mental health; psychotherapeutic processes; return to work; occupational stress; health plan implementation

Introduction

Occupational Stress and Work-Related Medical Rehabilitation

Work-related stress, as observed in one-third of the German population [1], is an important risk factor for common mental disorders [2]. Conversely, mental disorders have become the leading reasons for long-term sickness absence [3] and premature pension in Germany [4]. The main purpose of inpatient medical rehabilitation is to restore or promote capacity to work of rehabilitants. Work-related medical rehabilitation (WMR) focusing on the workplace situation has been implemented successfully but needs to be complemented by work-related aftercare interventions to support rehabilitants during vocational reintegration [5]. However, evidence has been mixed. First, studies with intensified work-related orthopedic rehabilitation aftercare have not proven to be effective compared to aftercare without a focus on work-related topics [6]. On the contrary, participation in a graded return to work program reduced the relative risk of permanent work disability by about 40% as well as the time of welfare benefits owing to sickness absence compared with matched controls [7]. Yet, participation in aftercare programs following inpatient medical rehabilitation is low because of long waiting times, the lack of local aftercare providers, or incompatibility with family or work commitments [8,9].

Web-based interventions have been shown to be effective for a broad range of mental disorders, for example, depression [10], anxiety disorders [11], pain [12], substance abuse [13], and also improved physical activity [14] or a healthy diet and weight reduction [15] or psychosocial support for patients with chronic diseases [16]. The majority of the German working population (nearly 90%) is on the Web using the internet daily (72%) [17] and a substantial part is using it as a frequent source of health information (38%) [18]. Thus, Web- and mobile-based interventions may close gaps in routine care and improve diagnostics and treatment in medical rehabilitation [19] as widely accessible and cost-effective interventions [20].

Web-Based Aftercare

Initial results of self-guided Web-based stress management interventions have been mixed. In a recent meta-analysis for Web-based interventions, moderate effect sizes were reported for overall 26 studies in reducing work-related stress (d=0.43, 95% CI 0.31-0.54) [21]. An internet- and mobile-based stress management program has proven effective in reducing perceived stress (d=0.96 posttreatment; d=0.65 6-month follow-up) compared with a waiting list control group and improved other relevant parameters of mental health, for example, depression, anxiety, and emotional exhaustion [22,23]. However, Web-based interventions when following inpatient medical rehabilitation are still rare [24] and interventions tested in randomized controlled trials have rarely been transferred into routine care.

We developed a Web-based transdiagnostic aftercare program (*GSA-Online*; *Gesund und Stressfrei am Arbeitsplatz* [*Healthy and stress-less at the workplace*]) that aimed at improving vocational reintegration of rehabilitants after long-term sickness absence. In our previous randomized controlled trial, *GSA-Online* had a statistically significant positive influence on the subjective prognosis of gainful employment (d=0.13 at the end of the intervention and d=0.20 at the follow-up 12 months after study inclusion). Furthermore we could show positive effects on affective mental distress (eg, d=0.25 for generalized anxiety or d=0.18 for depressive symptoms at follow-up) of the rehabilitants participating in the intervention [25] compared to the participants of an active control group.

Study Aims and Research Question

The aims of this study were (1) to examine whether the further developed *GSA-Online plus* can be implemented into routine care of inpatient medical rehabilitation, (2) to identify characteristics of patients who have received the recommendation for *GSA-Online plus*, and (3) to determine perceived helpfulness of the intervention and satisfaction of the participants. With regard to our primary outcome, we hypothesized that (1) *GSA-Online plus* would be recommended as often as other established aftercare programs for rehabilitants with a special need for WMR-like psychological treatment, rehabilitation aftercare, or vocational rehabilitation; (2) at least 66% of the rehabilitants with a recommendation for *GSA-Online plus* would participate and write at least 1 diary entry after rehabilitation; and (3) completers, that is, regular users with at least 6 diary entries, would achieve (a) a more positive subjective prognosis of gainful employment, (b) an increased quality of life and lower perceived stress, and (c) an improvement of emotional distress such as depression and anxiety at the end of the intervention compared to discharge from rehabilitation.

Methods

Study Design

The study was conducted in 2 rehabilitation centers treating 3 different medical indications (psychosomatic and orthopedic rheumatic diseases). Inpatient medical rehabilitation entails a multimodal group-oriented approach, supplemented by individual therapy, addressing work ability of rehabilitants with chronic (>6 months), somatic, or psychological impairments [26]. In a pre-post design, data collection took place at 14 time points: after discharge from inpatient rehabilitation, that is, at the beginning (T1), once a week during (T2-T13; only active participants), and at the end of the aftercare intervention (T14).

Recruitment of the study was conducted for 10 months between June 2016 and March 2017. Inclusion criteria were (1) employment and a plan to return to work within 4 weeks after inpatient medical rehabilitation, (2) the ability to write in German language, (3) age between 18 and 59 years, and (4) a

XSL•FO
RenderX

private internet access. During inpatient rehabilitation, the treating physician could prescribe the self-explanatory program by handing out an information leaflet to the patient and document it as a recommendation [27]. The recommendation was collected by the study assistant in each clinic who gave further study information, collected written consent, and gave access data for *GSA-Online plus* to the participant.

Intervention

The intervention *GSA-Online plus* aimed to support occupationally stressed rehabilitants during their return to work. In a structured psychodynamic format [28], participants were instructed to identify and articulate interpersonal and intrapsychic problems during return to work. Participants got weekly personalized writing impulses from a trained and supervised therapist to help them write in the form of a diary about their experiences of returning to their workplace. Therapeutic commentaries (usually within 24 hours) to their Web-based diary entry processed interpersonal and intrapsychic problems and helped participants deal with their individual job-related problems and stabilize their working capacity. Usually, it took 20 to 40 min per week per participant for the therapist to write an answer to the diary entry and add another writing impulse. In addition to the previous version (*GSA-Online*), educational video clips were used to familiarize participants with the program and its features (see Multimedia Appendices 1-4). For a detailed description of the therapeutic rationale, see Beutel et al [28].

Measures

As primary outcomes, the medical referral rate was documented (frequency of recommendations of *GSA-Online plus*) and participants' utilization of *GSA-Online plus* was tracked with PIWIK (now Matomo [29]), a secure open Web analytics platform and assessed with self-constructed single items (eg, "Please indicate how often you have used *GSA-Online plus* since the end of your inpatient treatment.").

In pre- and postmeasurements, the secondary outcome subjective prognosis of gainful employment was assessed with the SPE-Scale [30] consisting of 3 items assessing a subjective rating of future employment until retirement age, the impairment of work ability by the current health status, as well as the plan to apply for a premature pension. The 3 items could be added up to a score between 0 and 3 with a higher score indicating a higher risk for premature pension. The capacity to work was assessed with the short form of the work ability index (WAI), a 7-item scale with a reliability of alpha=.78 in a German population (eg, "How do you estimate your current work ability in terms of physical work requirements?") [31]. Mental disorders were assessed with different subscales of the German version of the Patient Health Questionnaire (PHQ-D) by Löwe et al [32]. Depressive symptoms were assessed with the 9-item scale PHQ-9 (eg, "Over the last two weeks, how often have you been bothered by little interest and pleasure in doing things?") [33], with an internal consistency of alpha=.89. Stress symptoms were assessed with the 10-item stress module PHQ-Stress (eg, "Over the last four weeks, how often have you been bothered by worries about health?" *)* [34,35]. Anxiety symptoms were assessed with the 7-item scale for General Anxiety Disorder (GAD-7; eg, "Over the last 2 weeks, how often have you been bothered by feeling nervous, anxious or on edge?" , alpha=.92) [36]. Somatoform symptoms were assessed with the 8 items of the Somatic Symptom Scale-8 (eg, "During the past 7 days, how much have you been bothered by stomach or bowel problems?", alpha=.81) [37] also based on the PHQ. Psychosocial stressors were measured with the 4-item short form of the Perceived Stress Scale (PSS-4; eg, "In the last month how often have you felt you were unable to control the important things in your life?", alpha=.60-.82) [38]. General functioning was measured with the 3-item Sheehan Disability Scale (eg, "To what extent do your symptoms impair your functioning in your social life?", alpha=.89) [39,40], where each item can be scored from 0 to 10 resulting in a global score from 0 (unimpaired) to 30 (highly impaired). Resources were assessed with the 3-item Oslo Social Support Scale (eg, "How many people are so close to you that you can count on them if you have great personal problems?", alpha=.60) [41], with the 4-item Brief Resilient Coping Scale (eg, "I look for creative ways to alter difficult situations", internal consistency r=0.76) [42], with the Loneliness Scale [43] and with the questions on life satisfaction, a scale that consists of 2 8-item modules (general life satisfaction and satisfaction with health, alpha=.82-89) [44]. Patient satisfaction was assessed with the 8-item Client Satisfaction Questionnaire (eg, "How would you rate the quality of the service you have received?", alpha=.93) [45]. In a weekly query, participants were asked with 2 items on a 5-point Likert scale (from 0=not at all to 4=very) about their satisfaction with *GSA-Online plus* ("How satisfied are you with the feedback of the online therapist?" and "How helpful was the feedback of the online therapist?") and rated their overall health condition with 1 item of the German version of the EuroQol Questionnaire (EQ-5D) [46] ("Your own health status today") as well as their current work ability with the first item from the WAI [31] also known as the Work Ability Score (WAS) [47], each on a Likert scale from 0 to 10. At the end of the aftercare program, participants were asked if and how much they would pay for *GSA-Online plus* with self-constructed items.

All questionnaires were assessed with the Web-based survey platform, SoSci Survey [48], except the weekly assessments that have been directly implemented in the platform of *GSA-Online plus*.

Statistical Analyses

All statistical analyses were done with IBM SPSS Statistics 23 [49]. Recommendation rates and utilization of *GSA-Online plus* were analyzed with cross-sectional analyses and participants of the Web-based aftercare were compared with the population of all rehabilitants treated during the recruitment period, using descriptive statistics (chi-square tests and *t* tests, Mann-Whitney U tests if the required assumptions for parametric testing, for example, homogeneity of variance, were violated). Pre-post changes were analyzed with per protocol data as secondary outcomes with longitudinal data analysis (*t* tests, rmANOVA, and descriptive statistics). To estimate treatment effects, Cohen *d* was calculated for the comparison of mean scores with *t* tests, Cramér V for chi-square tests and the effect size r was calculated for the comparison of median scores with the Mann-Whitney U test.

XSL•FO
RenderX

For the weekly assessment of the general health status and the subjectively rated ability to work, missing data were replaced by the last observation carried forward (LOCF) procedure. To analyze improvement across time, a repeated measures analysis of variance was conducted.

Ethics Approval

The study protocol was approved by the Ethics Committee of the Federal State of the Rhineland Palatinate (Approval Number: 837.175.16(10494)).

Results

Recommendation Rates for *GSA-Online plus*

Of the 2562 rehabilitants in 2 rehabilitation centers, 112 (4.4%) got a recommendation for *GSA-Online plus* during inpatient rehabilitation, which was significantly lower than the referral rate to other face-to face aftercare interventions (Table 1). Recommendation rates were higher in psychosomatic rehabilitation, where psychological treatments were routinely recommended in 88% (1030/1147) versus 8% (117/1147) after orthopedic rehabilitation. Rehabilitation aftercare was more frequently prescribed in orthopedic (500/1389, 36%) than in psychosomatic rehabilitation (89/1172, 8%). A total of 11% (291/2561) received recommendations of vocational rehabilitation.

Table 1. Recommendation rates at discharge for aftercare in total and in the 2 rehabilitation centers.

Aftercare recommendation	Psychosomatic rehabilitation center (n=1172), n (%)	Orthopedic and rheumatoid rehabilitation center (n=1389), n (%)	Total (N=2561), n (%)	χ^2_1	*P* value	*d*	*Cramér V*
GSA-Online plus	69 (5.9)	43 (3.1)	112 (4.4)	11.845	.001	0.14	−0.068
Psychological Treatment	1030 (87.9)	117 (8.4)	1147 (44.8)	1623.45	<.001	2.63	−0.796
Rehabilitation aftercare	89 (7.6)	500 (36.0)	589 (23.0)	289.57	<.001	0.71	0.336
Vocational rehabilitation	123 (10.5)	168 (12.1)	291 (11.4)	1.616	.20	0.05	0.025

Patients' Characteristics Associated With *GSA-Online plus* Recommendation

Rehabilitants who were recommended *GSA-Online plus* were younger (mean 47.67 [SD 9.97] vs mean 50.31 [SD 9.5], t_{2559}=2.865, *P*=.004, *d*=0.28) and had longer rehabilitation treatments (median 35 days vs median 28 days, *U*=182642.00, *P*<.001, *r*=−0.12). Furthermore, they reported longer work disability before inpatient rehabilitation (*P*=.004), and were more often (marginally significant, *P*=.07) considered able to work at discharge than rehabilitants who did not receive a recommendation (Table 2).

Of the 112 rehabilitants with a recommendation for *GSA-Online plus*, 77% (86/112) gave written informed consent and registered on the Web to participate in *GSA-Online plus*. From these, 58.1% (50/86) wrote at least 6 (completers) and 41.9% (36/86) less than 6 diary entries (dropouts). Multimedia Appendix 5 shows the distribution of the number of diary entries. Dropouts wrote an average of almost 3 diary entries (mean 1.53, [SD 1.72]), completers wrote 10 diary entries (mean 10.34 [SD 2.00]; t_{84} =−21.4, *P*<.001, *d*=−4.68). Two-thirds of completers (66.0%) wrote 11 or 12 diary entries.

Completers were older (mean 49.36 [SD 9.04] vs mean 45.14 [SD 10.49], t_{84}=−1.998, *P*=.049, *d*=0.44) and more often female (74.5% vs 50%, χ^2_1=4.638, *P*=.003) than dropouts. Further sociodemographic variables revealed no significant differences between completers and dropouts.

XSL•FO
RenderX

Table 2. Sociodemographic characteristics (T1: baseline): comparison of rehabilitants with and without recommendation for *GSA-Online plus*.

Sociodemographic characteristics	With recommendation (n=112), n (%)	Without recommendation (n=2449), n (%)	Total (N=2561), n (%)
Sex[a]			
Male	45 (40.2)	1067 (43.6)	1112 (43.4)
Female	67 (59.8)	1382 (56.4)	1449 (56.6)
Marital status[b]			
Partnership	60 (53.6)	1525 (62.3)	1585 (61.9)
No partnership	47 (42.0)	768 (31.4)	815 (31.8)
Unknown	5 (4.5)	156 (6.4)	161 (6.3)
Inability to work before inpatient admission[c]			
No inability or not employed	11 (9.8)	347 (14.2)	358 (14.0)
3 months or less	52 (46.4)	1061 (43.3)	1113 (43.5)
3 to 6 months	29 (25.9)	370 (15.1)	399 (15.6)
6 months or more	20 (17.9)	671 (27.4)	691 (27.0)
Work ability at discharge[d]			
Able to work	67 (59.8)	1203 (49.1)	1270 (49.6)
Unable to work	44 (39.3)	1196 (48.8)	1240 (48.4)
Unknown	1 (0.9)	50 (2.0)	51 (2.0)

[a]χ^2_1=0.501, *P*=.50, *Cramér V*=0.014.

[b]χ^2_2=5.713, *P*=.06, *Cramérs V*=0.047.

[c]χ^2_3=13.295, *P*=.004, *Cramérs V*=0.072.

[d]χ^2_2=5.20, *P*=.07, *Cramérs V*=0.045.

Effectiveness of *GSA-Online plus* and Participants' Satisfaction

In Figure 1 the weekly ratings of work ability and subjective health status are displayed. Missing data were replaced with the last observation carried forward method.

A statistically significant improvement could be observed for the subjective rated work ability assessed with the WAS ($F_{11,759}$=3.808, *P*<.001) but not for the subjective health status assessed with 1 item of the EQ-5D ($F_{11,759}$=3.928, *P*=.22).

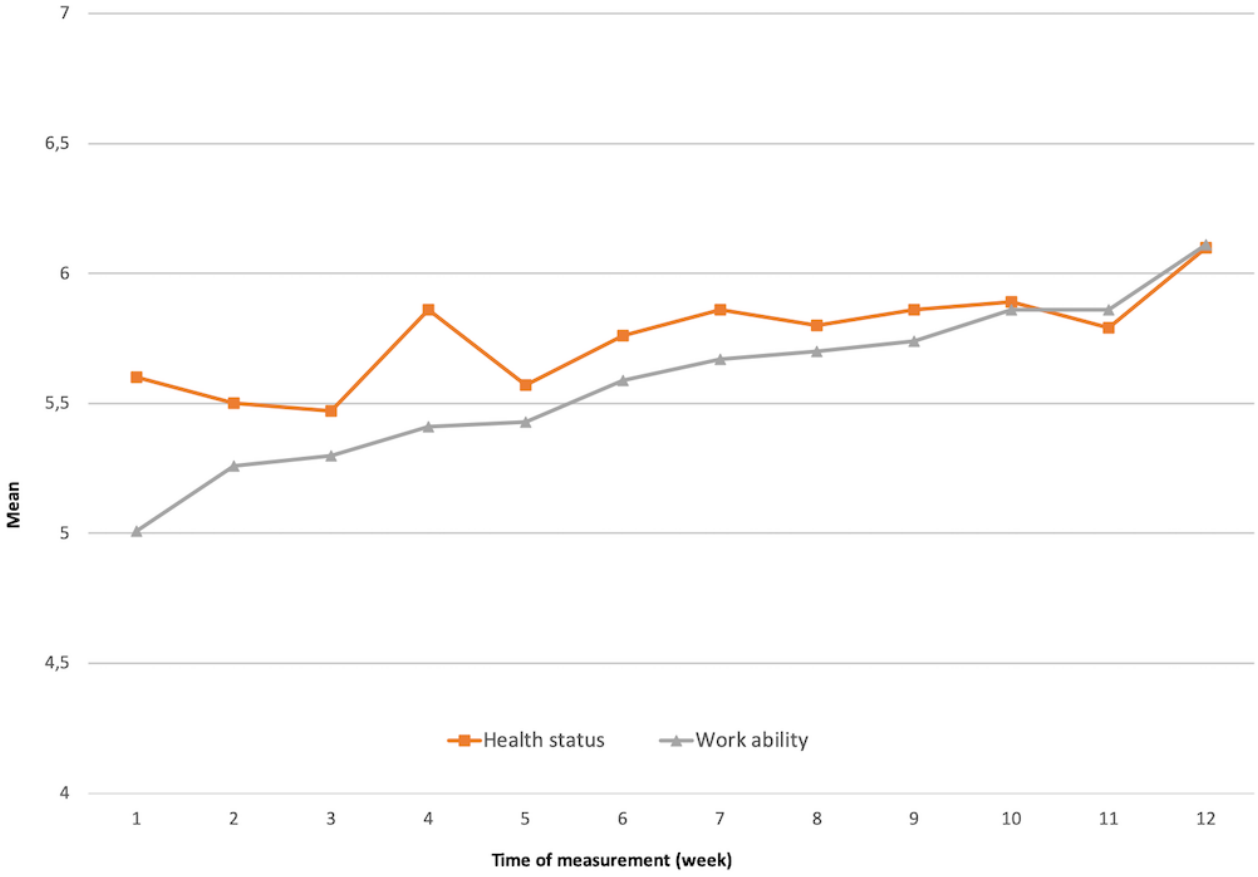
Satisfaction with *GSA-Online plus* was generally high. 43.5% (20/46) were largely and 45.7% (21/46) were very satisfied with *GSA-Online plus* overall and the level of support was also rated very positive. In total, 47.8% (22/46) were largely and 32.6% (15/46) were very satisfied with the help they got.

As displayed in Multimedia Appendix 6, for the weekly monitoring of satisfaction and rated helpfulness of the therapeutic feedback, a repeated measures analysis of variance was calculated. A statistically significant improvement (Multimedia Appendix 6) could be observed for the satisfaction with the feedback from a Web-based therapist ($F_{11,176}$=2.005, *P*=.03) as well as for the subjectively rated helpfulness ($F_{11,176}$=2.919, *P*=.001)—both assessed weekly with 1 item on a Likert scale from 0 to 4. For the purpose of sensitivity analysis, data of participants (N=17) who wrote 12 diary entries were analyzed separately leading to a clear increase of satisfaction and helpfulness (Multimedia Appendix 7).

Table 3 shows secondary outcome measures for regular users of *GSA-Online plus* (completers). Statistically significant improvements could be observed between discharge from the rehabilitation center and the end of the aftercare program (ie, 3 months after discharge) for subjective work ability as assessed by the WAI (*P*=.02, *d*=0.41), perceived stress assessed with the PSS-4 scale (*P*=.01, *d*=−0.38), functioning assessed with the SDS (*P*=.002, *d*=−0.60), and life satisfaction (*P*=.008, *d*=0.42). The other symptom scales scores remained stable and did not change between discharge and the follow-up assessment.

XSL•FO
RenderX

Figure 1. Means of the subjective health status and subjective work ability, assessed in the course of the aftercare from T2 to T13. Subjective health status was assessed with the EuroQoL-5D single item "How good or bad is your health status today?" 0=worst imaginable health; 10=best imaginable health. Subjective work ability was assessed with the Work Ability Score "Current ability to work in comparison with the best, ever reached ability to work"; (range 0-10). N=70, missing data replaced with last observation carried forward.



XSL•FO
RenderX

Table 3. Secondary outcome measures of completers of *GSA-Online plus* at discharge from inpatient medical rehabilitation (T1) and at the end of aftercare (T14).

Secondary outcomes	T1[a], mean (SD)	T14[b], mean (SD)	*P* value[c]	Effect size (*d*)[d]
Prognosis of gainful employment (n=44)	0.909 (1.007)	0.932 (0.998)	.82	0.03
Work ability (n=44)	25.080 (5.968)	27.125 (7.030)	*.02*[e]	0.41
Depression (n=46)	9.565 (4.420)	8.870 (4.400)	.17	−0.21
Anxiety (n=46)	7.348 (4.132)	7.587 (4.203)	.66	0.07
Somatic symptom (n=45)	11.667 (6.142)	11.733 (5.618)	.90	0.02
Psychosocial stress (n=46)	7.396 (3.655)	7.913 (3.601)	.26	0.17
Perceived stress (n=44)	7.955 (1.976)	7.136 (2.339)	*.01* [e]	−0.38
General functioning (n=44)	22.386 (5.406)	18.318 (7.398)	*.002* [f]	−0.60
Loneliness (n=44)	7.727 (3.022)	7.144 (3.166)	.46	−0.27
Social support (n=44)	9.272 (2.386)	9.386 (2.315)	.60	0.08
Coping (n=43)	14.047 (3.280)	14.907 (3.069)	.09	0.16
Life satisfaction (n=43)	25.581 (5.261)	26.814 (5.193)	*.008* [f]	0.42

[a]Baseline.

[b]End of intervention.

[c]*t*-test for dependent samples; statistically significant differences are italicized.

[d]*d*: Cohen *d*.

[e]Significant at $P<.05$.

[f]Significant at $P<.01$.

Of the 46 participants who completed the Web-based questionnaire at measurement T14, the majority (27/46, 58.7%) watched at least one of the educational video clips but also almost half (19/46, 41.3%) had not seen any of the films. The main reasons for not watching any films were that the existence of the films was unknown to the participants (8/19, 42%), participants did not use any other features of *GSA-Online plus* besides the Web-based diary (5/19, 26%) or reported a lack of time (5/19, 26%). In total, 77% (23/30) of the participants who answered questions about video use assessed the films as a positive contribution to the comprehensibility of *GSA-Online plus*.

Finally, the willingness of participants to pay for *GSA Online Plus* was assessed at T14. Participants were asked if they would be willing to pay *for GSA-Online plus* and how much they would pay if they were required to pay. Of the 46 participants who answered the first question, 37% (17/46) said they were willing to pay for *GSA-Online plus* and 40 participants reported to pay in average (mean) 174.25 Euro (SD 292.2, Min 0, Max 1500).

Discussion

Principal Findings

Inpatient medical rehabilitation serves to promote work ability and vocational reintegration is a crucial outcome. Although previous Web-based trials have improved coping with work-related stress [23] or showed that Web-based aftercare could successfully maintain effects of inpatient treatment [50], information on implementation processes are missing because previous Web-based interventions have been conducted under study conditions, and participants were usually recruited on the Web. The purpose of the trial was to evaluate the implementation of a Web-based aftercare program (*GSA-Online plus*) during inpatient medical rehabilitation.

As we had hypothesized, it was feasible to implement the program, and the great majority who received the recommendation (86/112, 77%) actually logged into the program, slightly exceeding our expectations. Among those who started writing blogs, subjective or perceived rated work ability increased over the course of their participation. Satisfaction with Web-based aftercare was generally high. Acceptance was good with 43.5% (20/46) of the participants being largely and 45.7% (21/46) very satisfied with *GSA-Online plus* overall and the level of support was also rated very positively with 47.8% (22/46) of the participants being largely and 32.6% (15/46) very satisfied with the help they received. Contrary to our hypotheses, subjective health status did not increase significantly. However, there was a significant increase of subjective work ability, general functioning, as well as life satisfaction and a decrease of subjective stress. Especially, work ability and functioning are major issues in rehabilitation; therefore, a stabilization of these factors is a good indicator that our aftercare intervention had a focus on the right topics. Unfortunately, no change was found regarding mental distress for the completers of the intervention with a per protocol analysis. An explanation could be, that mental distress was not so high in this sample, with not only psychiatric but also orthopedic main diagnoses, therefore, an improvement was hard to detect in the relatively small sample we could include in our pre-post analyses.

XSL•FO
RenderX

However, unexpectedly, the referral rate by the physicians of about 4% was substantially lower than all face-to-face aftercare offers and there was also a lot of individual variation among physicians. Recommendations were made more frequently in psychosomatic (5.9% [69/1172]) than orthopedic rehabilitation (3.1% [43/1389]) and more often to younger patients with extended sick leaves and lengthy inpatient stays. This finding coincides with the results of Hennemann et al [51], showing clear preconceptions and barriers regarding Web-based contact among rehabilitation staff, concerning disruption of face-to-face therapeutic alliance and a lack of sufficient data security in Web-based aftercare interventions.

Owing to the low proportion of recommendations, we asked members of the treatment team of the rehabilitation centers in a Web-based survey at the end of the study for their subjective criteria and frequencies of recommendation of *GSA-Online plus*. Unfortunately, only a few members (N=19) of the treatment team took part in this assessment and the biggest part (8/19, 61.5%) made less than 10 referrals for *GSA-Online plus*, whereas 23.1% (3/19) referred 11 to 20 and 15.4% (2/19) more than 20 rehabilitants. Members of the treatment team were informed about *GSA-Online plus* in a one-time on-site training and an informational paper that was accessible to all throughout the study. One explanation for a different referral rate between different team members could well be that especially those who did not attend the training did not make so many recommendations. Data we have on this point at least suggest that there was a trend that referral rate of those who did not attend the training was somehow lower than of those who attended the training. Asked, whether Web-based interventions could be a useful supplement or substitute for regular face-to-face interventions the majority (8/19, 42.1%) answered that they could be a useful supplement but no one rated them as a substitute to usual care.

Our finding is of special interest, because the 121st German congress of physicians has recently lifted the ban on remote treatment [52]. Furthermore, the German statutory pension insurance scheme recently developed requirements for the implementation of tele medical aftercare programs into routine care [53].

Previously, reservations towards occupational e-mental health interventions were also found in patients, especially in the risk groups the interventions were planned for [54]. In spite of the low recommendation rate, the intervention has shown promise in participants. Thus, physicians are in a prime position to provide access to and motivate their patients for sustained participation in these programs. Our findings implicate that 2 steps should be pursued in the future: (1) Awareness about the benefits of Web-based interventions in routine care and their compatibility as an adjunct to face-to face aftercare are required to promote openness in the treatment team to recommend, and to increase referral rates for innovative interventions. (2) Physicians need to be instructed about their patients' reservations to sign up for Web-based support and enabled to help their patients initiate and maintain their engagement in suitable Web-based programs.

Limitations

Limitations of our study were that we had no control group to analyze efficacy in a routine care setting and that we did not reach our anticipated sample size, because of the much lower referral rate than expected. It also may be debated if logging in once is a sufficient criterion for participation.

The LOCF method for imputing missing data is not the gold standard but seemed sufficient in this small sample, especially as efficacy only was a secondary research question. Furthermore, the multiple comparisons within our secondary research questions are to be considered as a limitation. After adjustment for multiple testing, only general functioning will remain as a significant improvement after participating in *GSA-Online plus*.

Unfortunately, we did not ask participants about advantages of Web-based interventions but we asked them about their digital competencies. Two-thirds of the participants rated their knowledge of using digital media (eg, PC, smartphone, and tablet) as mediocre to rather good. Overall, the attitude toward the internet was positive, but one-third also stated that they feel burdened by the constant availability via mobile phone or email.

Hence, future implementations should focus more on collaborating with staff and clinicians in rehabilitation clinics to address potential prejudices and barriers to Web-based aftercare. On the patient side, it is important to address advantages and disadvantages of Web-based interventions to improve acceptance for internet- and mobile-based interventions.

Conclusions

Web-based psychological aftercare proved to be effective to reduce treatment gaps after inpatient medical rehabilitation but only for a limited number of rehabilitants. *GSA-Online plus* has been provided as a Web-based aftercare that offers all inpatient medical rehabilitation patients with occupational exposure, mental comorbidity, and an intentional timely return to work the opportunity to be promoted and assisted with professional reintegration by trained psychologists. Once a recommendation for *GSA-Online plus* was given from the physician in routine care, it could lead to a significantly higher participant motivation and adherence than in a controlled efficacy study. Outcome criteria on which the Web-based aftercare focuses (ability to function, work ability, general state of health, and life satisfaction) will improve even further in the course of follow-up, compared with the state of health at the end of inpatient rehabilitation. Motivation of rehabilitants and attitudes of the treatment team toward Web-based interventions are essential to improve implementation/recommendation rates. An important question for future research could be how Web-based interventions for rehabilitants with work-related problems could be optimized or supplemented to reach more rehabilitants. As the return to work is a major issue in rehabilitative treatment in Germany, a more practical oriented intervention with a more social work-driven focus could possibly close this gap and continue the multidisciplinary treatment approach that is already one main characteristic of inpatient medical rehabilitation in Germany.

XSL•FO
RenderX

Acknowledgments

The authors would like to express their gratitude to all participants of the trial as well as the research assistants of the Mittelrhein-Klinik in Boppard/Bad Salzig, Germany (Melanie Mayeh, Dietmar Kraeber), the Karl Aschoff Klinik in Bad Kreuznach, Germany (Anja Zocher-Tiegs) and the Drei-Burgen-Klinik, Bad Kreuznach (Michael Wiegert, Vera Simon-Reichert). This study was funded by the German Statutory Pension Insurance Scheme (Deutsche Rentenversicherung Bund; grant number: 0423/40–64–50-50) from 08/2015 to 10/2017.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot of the landing page.

[PDF File (Adobe PDF File), 690KB - jmir_v21i6e12285_app1.pdf]

Multimedia Appendix 2

Overview of the online diary.

[PDF File (Adobe PDF File), 203KB - jmir_v21i6e12285_app2.pdf]

Multimedia Appendix 3

Screenshot of an explanatory video.

[PDF File (Adobe PDF File), 1MB - jmir_v21i6e12285_app3.pdf]

Multimedia Appendix 4

Screenshot of an explanatory video.

[PDF File (Adobe PDF File), 1MB - jmir_v21i6e12285_app4.pdf]

Multimedia Appendix 5

Number of written online diary entries (completer vs dropouts).

[PDF File (Adobe PDF File), 34KB - jmir_v21i6e12285_app5.pdf]

Multimedia Appendix 6

Means of the weekly monitoring of satisfaction and rated helpfulness of the therapeutic feedback, both assessed weekly with one item each on a Likert-scale from 0 to 4. N=70, missing data replaced with last observation carried forward.

[PDF File (Adobe PDF File), 58KB - jmir_v21i6e12285_app6.pdf]

Multimedia Appendix 7

Means of the weekly monitoring of satisfaction and rated helpfulness of the therapeutic feedback, both assessed weekly with one item each on a Likert-scale from 0 to 4. N=17, per protocol data.

[PDF File (Adobe PDF File), 62KB - jmir_v21i6e12285_app7.pdf]

References

1. Bethge M, Radoschewski FM, Müller-Fahrnow W. Work stress and work ability: cross-sectional findings from the German sociomedical panel of employees. Disabil Rehabil 2009 Sep;31(20):1692-1699. [doi: 10.1080/09638280902751949] [Medline: 19479539]
2. Stansfeld S, Candy B. Psychosocial work environment and mental health—a meta-analytic review. Scand J Work Environ Health 2006 Dec;32(6):443-462. [doi: 10.5271/sjweh.1050]
3. Henderson M, Harvey SB, Overland S, Mykletun A, Hotopf M. Work and common psychiatric disorders. J R Soc Med 2011 May;104(5):198-207 [FREE Full text] [doi: 10.1258/jrsm.2011.100231] [Medline: 21558098]

XSL•FO
RenderX

4. [German Pension Insurance]. Berlin: Deutsche Rentenversicherung Bund; 2014. [Position paper of the Deutsche Rentenversicherung on the importance of mental illness in rehabilitation and reduction of earning capacity] URL: https://tinyurl.com/y2uvz9ql [accessed 2019-06-04]
5. Bethge M. [Work-related medical rehabilitation]. Rehabilitation (Stuttg) 2017 Feb;56(1):14-21. [doi: 10.1055/s-0042-118579] [Medline: 28219096]
6. Briest J, Bethge M. [Intensified work-related rehabilitation aftercare: long-term results of a randomized controlled multicenter trial]. Rehabilitation (Stuttg) 2016 Apr;55(2):108-114. [doi: 10.1055/s-0042-102998] [Medline: 27070985]
7. Bethge M. Effects of graded return-to-work: a propensity-score-matched analysis. Scand J Work Environ Health 2016 Dec 1;42(4):273-279 [FREE Full text] [doi: 10.5271/sjweh.3562] [Medline: 27074050]
8. Kobelt A, Nickel L, Grosch E, Lamprecht F, Künsebeck HW. [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. 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 (Stuttg) 2011 Dec;50(6):363-371. [doi: 10.1055/s-0031-1271815] [Medline: 21647850]
10. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
11. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
12. Macea DD, Gajos K, Daglia CY, Fregni F. The efficacy of web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
13. Rooke S, Thorsteinsson E, Karpin A, Copeland J, Allsop D. Computer-delivered interventions for alcohol and tobacco use: a meta-analysis. Addiction 2010 Aug;105(8):1381-1390. [doi: 10.1111/j.1360-0443.2010.02975.x] [Medline: 20528806]
14. Davies C, Spence J, Vandelanotte C, Caperchione C, Mummery W. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
15. Grunenberg E, Lin J, Baumeister H. [Effectiveness of web-based psychological interventions for weight loss--a systematic review]. Rehabilitation (Stuttg) 2013 Jun;52(3):182-187. [doi: 10.1055/s-0033-1343489] [Medline: 23761207]
16. Kanera IM, Willems RA, Bolman CA, Mesters I, Zambon V, Gijsen BC, et al. Use and appreciation of a tailored self-management eHealth intervention for early cancer survivors: process evaluation of a randomized controlled trial. J Med Internet Res 2016;18(8):e229 [FREE Full text] [doi: 10.2196/jmir.5975] [Medline: 27554525]
17. Koch W, Frees B. [ARD-ZDF Online Study | ARD / ZDF Media Commission]. ARD/ZDF-Onlinestudie 2017: Neun von zehn Deutschen online. Media perspektiven. 2017; 2017. [ARD/ZDF online study 2017: Nine out of ten Germans online] URL: http://www.ard-zdf-onlinestudie.de/files/2017/Artikel/917_Koch_Frees.pdf [accessed 2019-06-06] [WebCite Cache ID 78vQqmO5m]
18. 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]
19. Baumeister H, Lin J, Ebert DD. [Internet- and mobile-based approaches : psycho-social diagnostics and treatment in medical rehabilitation]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2017 Apr;60(4):436-444. [doi: 10.1007/s00103-017-2518-9] [Medline: 28224187]
20. Donker T, Blankers M, Hedman E, Ljótsson B, Petrie K, Christensen H. Economic evaluations of Internet interventions for mental health: a systematic review. Psychol Med 2015 Dec;45(16):3357-3376. [doi: 10.1017/S0033291715001427] [Medline: 26235445]
21. Heber E, Ebert DD, Lehr D, Cuijpers P, Berking M, Nobis S, et al. The benefit of web-and computer-based interventions for stress: a systematic review and meta-analysis. J Med Internet Res 2017 Feb 17;19(2):e32 [FREE Full text] [doi: 10.2196/jmir.5774] [Medline: 28213341]
22. Ebert DD, Heber E, Berking M, Riper H, Cuijpers P, Funk B, et al. Self-guided internet-based and mobile-based stress management for employees: results of a randomised controlled trial. Occup Environ Med 2016 May;73(5):315-323. [doi: 10.1136/oemed-2015-103269] [Medline: 26884049]
23. Heber E, Lehr D, Ebert DD, Berking M, Riper H. Web-based and mobile stress management intervention for employees: a randomized controlled trial. J Med Internet Res 2016;18(1):e21 [FREE Full text] [doi: 10.2196/jmir.5112] [Medline: 26818683]
24. Eichenberg C, Ott R. [Clinical-psychological intervention on the internet]. Psychotherapeut 2011 Jun 10;57(1):58-69 [FREE Full text] [doi: 10.1007/s00278-011-0832-5]
25. Zwerenz R, Becker J, Gerzymisch K, Siepmann M, Holme M, Kiwus U, et al. Evaluation of a transdiagnostic psychodynamic online intervention to support return to work: A randomized controlled trial. PLoS One 2017;12(5):e0176513 [FREE Full text] [doi: 10.1371/journal.pone.0176513] [Medline: 28481893]

XSL•FO
RenderX

26. Linden M. Psychosomatic inpatient rehabilitation: the German model. Psychother Psychosom 2014;83(4):205-212. [doi: 10.1159/000358852] [Medline: 24970244]
27. Zwerenz R, Böhme K, Wirth A, Labitzke N, Pachtchenko S, Beutel ME. Integration of the work-related online aftercare intervention 'GSA-online plus' (healthy and without stress at the workplace) into clinical practice: study protocol for an implementation study. BMC Health Serv Res 2018 Dec 2;18(1):312 [FREE Full text] [doi: 10.1186/s12913-018-2995-z] [Medline: 29716605]
28. Beutel M, Böhme K, Banerjee M, Zwerenz R. Psychodynamic online treatment following supportive expressive therapy (SET): therapeutic rationale, interventions and treatment process. Z Psychosom Med Psychother 2018 Jun;64(2):186-197. [doi: 10.13109/zptm.2018.64.2.186] [Medline: 29862925]
29. Matomo: #1 Secure Open Web Analytics Platform. URL: https://matomo.org [accessed 2019-06-04] [WebCite Cache ID 78vYkRjXk]
30. 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 (Stuttg) 2003 Jun;42(3):169-174. [doi: 10.1055/s-2003-40095] [Medline: 12813654]
31. Ilmarinen J. The Work Ability Index (WAI). Occup Med (Lond) 2006 Oct 17;57(2):160. [doi: 10.1093/occmed/kqm008]
32. Löwe B, Spitzer R, Zipfel S, Herzog W. [University Hospital Heidelberg - University of Heidelberg]. Karlsruhe: Pfizer; 2002. [PHQ-D: Health questionnaire for patients] URL: https://www.klinikum.uni-heidelberg.de/fileadmin/Psychosomatische_Klinik/download/PHQ_Manual1.pdf [accessed 2019-06-04] [WebCite Cache ID 78s9N4k7g]
33. 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]
34. Gräfe K, Zipfel S, Herzog W, Löwe B. [Screening of mental disorders with the "Health Questionnaire for Patients (PHQ-D)"]. Diagnostica 2004 Oct;50(4):171-181. [doi: 10.1026/0012-1924.50.4.171]
35. Klapow J, Kroenke K, Horton T, Schmidt S, Spitzer RB, Williams JB. Psychological disorders and distress in older primary care patients: a comparison of older and younger samples. Psychosom Med 2002;64(4):635-643. [doi: 10.1097/01.psy.0000021942.35402.c3] [Medline: 12140354]
36. 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]
37. Gierk B, Kohlmann S, Kroenke K, Spangenberg L, Zenger M, Brähler E, et al. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. JAMA Intern Med 2014 Mar;174(3):399-407. [doi: 10.1001/jamainternmed.2013.12179] [Medline: 24276929]
38. Cohen S, Williamson G. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Social Psychology of Health: The Claremont Symposium on Applied Social Psychology. Newbury Park, CA: Sage Publications; 1988:31-67.
39. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol 1996 Jun;11(Suppl 3):89-95. [Medline: 8923116]
40. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. Int J Psychiatry Med 1997;27(2):93-105. [doi: 10.2190/T8EM-C8YH-373N-1UWD] [Medline: 9565717]
41. Dalgard OS. Community mental health profile as tool for psychiatric prevention. In: Trent DR, Reed C, editors. Promotion of Mental Health, 1995, Volume 5. Avebury: Ashgate Publishing, Limited; 1996.
42. Kocalevent R, Zenger M, Hinz A, Klapp B, Brähler E. Resilient coping in the general population: standardization of the brief resilient coping scale (BRCS). Health Qual Life Outcomes 2017 Dec 28;15(1):251 [FREE Full text] [doi: 10.1186/s12955-017-0822-6] [Medline: 29282066]
43. Richter D, Weinhardt M. LS-S: Loneliness Scale-SOEP. In: Kemper CJ, Zenger M, Brähler E, editors. [Psychological and Social Sciences Short Scales: Standardized Survey Instruments for Science and Practice]. Berlin: MWV Medizinisch Wissenschaftliche Verlagsgesellschaft; 2014:192-193.
44. Henrich G, Herschbach P. Questions on Life Satisfaction (FLZM): a short questionnaire for assessing subjective quality of life. Eur J Psychol Assess 2000 Sep;16(3):150-159. [doi: 10.1027//1015-5759.16.3.150]
45. Schmidt J, Lamprecht F, Wittmann WW. [Satisfaction with inpatient care: development of a questionnaire and first validity assessments]. Psychother Med Psychol 1989;39(7):248-255.
46. Graf J, Claes C, Greiner W, Uber A. [The German version of the EuroQol questionnaire]. J Public Health 1998 Mar;6(1):3-20. [doi: 10.1007/BF02956350]
47. El Fassi M, Bocquet V, Majery N, Lair ML, Couffignal S, Mairiaux P. Work ability assessment in a worker population: comparison and determinants of Work Ability Index and Work Ability score. BMC Public Health 2013 Apr 8;13:305 [FREE Full text] [doi: 10.1186/1471-2458-13-305] [Medline: 23565883]
48. Leiner D. SoSci Survey. SoSci Survey (Version 3.1.06) URL: https://www.soscisurvey.de [accessed 2019-06-04] [WebCite Cache ID 78sB7xiya]
49. IBM Corp. Armonk, NY: IBM Corp; 2014. IBM SPSS Statistics for Windows, Version 23 URL: https://www-01.ibm.com/support/docview.wss?uid=swg24038592 [accessed 2019-06-06] [WebCite Cache ID 78vU8qlZB]

XSL•FO
RenderX

50. 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]
51. Hennemann S, Beutel ME, Zwerenz R. Ready for eHealth? Health professionals' acceptance and adoption of eHealth interventions in inpatient routine care. J Health Commun 2017 Mar;22(3):274-284. [doi: 10.1080/10810730.2017.1284286] [Medline: 28248626]
52. Korzilius H. [Opening of the 121th German Medical Conference: wrestling for correct arguments]. Dtsch Arztebl Int 2018;115(19):907-909 [FREE Full text]
53. Deutsche Rentenversicherung Bund. Anforderungen an Tele-Reha-Nachsorge. Berlin: Deutsche Rentenversicherung Bund; 2018. URL: https://tinyurl.com/y6jr4y5b [accessed 2019-06-04]
54. Hennemann S, Witthöft M, Bethge M, Spanier K, Beutel ME, Zwerenz R. Acceptance and barriers to access of occupational e-mental health: cross-sectional findings from a health-risk population of employees. Int Arch Occup Environ Health 2018 Apr;91(3):305-316. [doi: 10.1007/s00420-017-1280-5] [Medline: 29189895]

Abbreviations

EQ-5D: EuroQol questionnaire
GSA: Gesund und Stressfrei am Arbeitsplatz [Healthy and stress-less at the workplace]
LOCF: Last observation carried forward
PHQ: Patient Health Questionnaire
PSS-4: Perceived Stress Scale
WAI: Work ability index
WAS: Work Ability Score
WMR: Work-related medical rehabilitation

Edited by G Eysenbach; submitted 02.10.18; peer-reviewed by E Neter, J Lin, C Eichenberg, K Spanier; comments to author 22.11.18; revised version received 20.02.19; accepted 26.04.19; published 18.06.19.

Please cite as:
Zwerenz R, Baumgarten C, Dahn I, Labitzke N, Schwarting A, Rudolph M, Ferdinand P, Dederichs-Masius U, Beutel ME
Implementation of a Web-Based Work-Related Psychological Aftercare Program Into Clinical Routine: Results of a Longitudinal Observational Study
J Med Internet Res 2019;21(6):e12285
URL: http://www.jmir.org/2019/6/e12285/
doi:10.2196/12285
PMID:31215515

©Rüdiger Zwerenz, Carlotta Baumgarten, Ingo Dahn, Nicole Labitzke, Andreas Schwarting, Matthias Rudolph, Peter Ferdinand, Ute Dederichs-Masius, Manfred E Beutel. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Comparing Treatment Acceptability and 12-Month Cessation Rates in Response to Web-Based Smoking Interventions Among Smokers Who Do and Do Not Screen Positive for Affective Disorders: Secondary Analysis

Noreen L Watson[1], PhD; Jaimee L Heffner[1], PhD; Kristin E Mull[1], MS; Jennifer B McClure[2], PhD; Jonathan B Bricker[1,3], PhD

[1]Fred Hutchinson Cancer Research Center, Seattle, WA, United States

[2]Kaiser Permanente Washington Health Research Institute, Seattle, WA, United States

[3]University of Washington, Seattle, WA, United States

Corresponding Author:
Noreen L Watson, PhD
Fred Hutchinson Cancer Research Center
1100 Fairview Ave N
M3-B232
Seattle, WA,
United States
Phone: 1 206 667 2942
Email: nlwatson@fredhutch.org

Abstract

Background: Web-based cessation programs are now common for intervening with smokers. However, it remains unclear how acceptable or effective these interventions are among people with affective disorders and symptoms (ADS; eg, depression and anxiety). Research examining this is extremely limited, with mixed results on cessation rates. Additional large studies are needed to more fully understand whether Web-based interventions are similarly used and equally effective among people with and without affective disorder symptomology. If not, more targeted Web-based interventions may be required.

Objective: The goal of the research was to compare Web-based treatment acceptability (defined by satisfaction and use) and 12-month cessation outcomes between smokers with and without ADS.

Methods: Participants (N=2512) were adult smokers enrolled in a randomized, comparative effectiveness trial of two Web-based smoking interventions designed for the general population of smokers. At baseline, participants reported demographic and smoking characteristics and completed measures assessing ADS. Participants were then classified into subgroups based on their self-reported ADS—either into a no ADS group or into six nonmutually exclusive subgroups: depression, posttraumatic stress disorder (PTSD), panic disorder (PD), generalized anxiety disorder (GAD), social anxiety disorder (SAD), and more than one ADS. Surveys at 12 months postrandomization included subjective ratings of treatment acceptability and self-reported smoking cessation. Treatment use (ie, number of log-ins and total duration of exposure) was assessed via automated records.

Results: Relative to the no ADS group, all six ADS subgroups reported significantly greater satisfaction with their assigned Web treatment program, but they spent less time logged in than those with no ADS. For number of log-ins, a treatment arm by ADS group interaction was observed across all ADS subgroups except GAD, suggesting that relative to the no ADS group, they logged in less to one website but not the other. At the 12-month follow-up, abstinence rates in the no ADS group (153/520, 29.42%) were significantly higher than for participants who screened positive for depression (306/1267, 24.15%; *P*=.03), PTSD (294/1215, 24.19%; *P*=.03), PD (229/1003, 23.83%; *P*=.009), and two or more ADS (323/1332, 24.25%; *P*=.03). Post hoc analyses suggest the lower quit rates may be associated with differences in baseline nicotine dependence and levels of commitment to resist smoking in difficult situations. Website use did not explain the differential abstinence rates.

Conclusions: Despite reporting higher levels of treatment satisfaction, most smokers with ADS used their assigned intervention less often and had lower quit rates than smokers with no ADS at treatment onset. The results support the need for developing more targeted interventions for smokers with ADS.

XSL•FO
RenderX

Trial registration: Clinical Trials.gov NCT01812278; https://clinicaltrials.gov/ct2/show/NCT01812278 (Archived by WebCite at http://www.webcitation.org/78L9cNdG4)

(J Med Internet Res 2019;21(6):e13500) doi:10.2196/13500

KEYWORDS

smoking; smoking cessation; affective disorders; anxiety; depression; eHealth; Web intervention; co-occurring disorders

Introduction

With lifetime prevalence rates ranging from 16.6% to 28.8%, affective disorders such as depression and anxiety are among the most common mental health conditions in the United States [1]. Epidemiological data indicate that individuals with these disorders, as well as those with elevated symptoms of anxiety or depression that do not meet a diagnostic threshold, have significantly higher prevalence rates of smoking and are less likely to successfully quit than those without such conditions or symptoms [2-6]. As a result, smokers with affective disorders and symptoms (ADS) account for disproportionate rates of tobacco-related deaths and diseases [7,8], and ongoing efforts are needed to improve cessation rates among this at-risk group of smokers.

One way to increase rates of cessation among smokers with ADS is to make treatment more accessible. This can be accomplished with interventions that are low-burden, low-cost, and high-reach, including Web-based interventions. Indeed, each year an estimated 12 million smokers look online for help to quit smoking [9]. In the general population of smokers, Web-based interventions demonstrated similar levels of effectiveness as more traditional treatment modalities such as face-to-face and telephone-delivered interventions in a recent meta-analysis [10]. However, there is a paucity of research examining the acceptability and effectiveness of these nontargeted interventions for individuals exhibiting ADS, who may have unique needs not addressed in Web-based interventions designed for the general population. Results from the few Web-based intervention studies reporting smoking cessation outcomes for individuals with ADS have yielded mixed findings with regard to their effectiveness for cessation. Of the two studies that examined anxiety, one demonstrated that individuals with greater levels of anxiety were less likely to quit and more likely to relapse compared with their less anxious counterparts at 4-week follow-up [11], whereas the other did not find a relationship between anxiety and quitting smoking at 3-month follow-up [12]. In addition, neither study examined how anxiety influenced long-term quit rates, both studies had low rates of outcome data retention, and both studies used broad measures of anxiety rather than comparing outcomes among different types of anxiety disorders or symptom categories. Thus, additional research addressing these limitations is needed. With regard to the impact of depression on smoking cessation when using Web-based interventions, two studies have demonstrated that 13-month quit rates were significantly lower among individuals with current depressive symptoms versus those with no depression [13,14]. However, three studies have demonstrated that current depressive symptoms were not associated with smoking cessation at 3 months [12,15] or 12 months [16]. In short, it remains unclear whether standard Web-based smoking interventions (ie, those not targeted for special populations) are less effective for individuals with ADS.

Based on research with other (non-Web) treatment modalities, one might expect that standard Web-based interventions would be less effective for smokers with ADS relative to those who do not have this symptomology [5,17-20] or at least less effective for some types of ADS. For example, because individuals with social anxiety disorder often avoid interactions with others due to fears about being negatively evaluated, socially anxious smokers may respond well to Web-based treatments, as they would provide access to an evidence-based intervention without needing to interact with other smokers or providers directly. On the other hand, smokers with affective disorders associated with various degrees of cognitive impairment (eg, poor concentration), such as depression or posttraumatic stress disorder (PTSD), may not respond as well to self-guided, stand-alone Web-based interventions. Examining such disorder-specific differentiations in smoking outcomes remains unexplored in the literature, but having a greater understanding how specific ADS and comorbid affective symptomology are associated with treatment satisfaction, use, and outcome could enhance future efforts to develop more effective interventions for these.

To best guide these targeted treatment development efforts, it is important to understand how smokers with different ADS respond differentially to Web-based cessation programs relative to smokers without ADS. Thus, the primary goal of this study was to compare treatment acceptability, defined both by self-reported satisfaction and treatment use, and 12-month abstinence rates among smokers with and without symptoms of one or more common affective disorders (depression, social anxiety disorder [SAD], generalized anxiety disorder [GAD], panic disorder [PD], and PTSD). All participants were recruited as part of a prior comparative effectiveness trial of two online interventions, each targeting the general population of smokers and testing two different Web-based interventions [21]. This paper is a secondary analysis of data collected from this randomized controlled trial (RCT). We hypothesized that, regardless of the assigned Web-based treatment program, smokers with ADS at treatment onset would have less satisfaction, lower levels of engagement, and lower cessation rates than smokers with no ADS and aimed to determine if there is evidence of differential outcomes that are disorder-specific.

Methods

Overview

A comprehensive description of study methods can be found in the main outcomes paper [21]. In brief, participants were

XSL•FO
RenderX

2512 of the 2637 adult smokers from all 50 states enrolled in the parent RCT, which compared the effectiveness of two Web-based smoking cessation programs [NCT01812278]. For the purposes of these analyses, the analytic sample was limited to participants who had complete data on at least one screening measure assessing affective disorder symptomology (see below for more details on these measures). To be eligible for the parent trial, participants had to (1) be 18 years of age or older; (2) smoke 5 or more cigarettes per day for the last year; (3) express a desire to quit smoking within 30 days; (4) have access to high-speed internet, including access to email; (5) not be participating in other smoking interventions or treatments; (6) never have used the control intervention, Smokefree.gov; (7) never have participated in one of our previous smoking studies; (8) not have another member in their household participating in the study; (9) express a willingness to be randomized to treatment, complete 3 follow-up surveys, and provide contact information for themselves and two relatives; (10) reside in the United States; and (11) be able to read in English. After completing the informed consent and baseline survey online, participants were randomized to one of two Web-based programs based either on acceptance and commitment therapy (ACT) or standard care (the National Cancer Institute's Smokefree.gov website). The ACT-based website presented content in a sequential manner through four phases and prompted users to track their smoking each day. The Smokefree.gov website had three main sections and users could navigate through all pages of the website freely with no restrictions on the order in which content had to be viewed. Participants had access to their assigned interventions for 12 months and received up to 4 messages per day for 28 days (via text or email) designed to encourage website use. Outcome surveys were completed at 3, 6, and 12 months postrandomization. The 12-month data retention rate was 88%, and neither treatment arm [21] nor mental health status predicted retention [22]. Further details on the recruitment and retention methods can be found elsewhere [22].

Measures

Baseline Demographics, Smoking, and Smoking-Related Characteristics

At baseline, participants reported demographic characteristics including age, gender, race and ethnicity, marital status, education, employment status, and sexual orientation. Level of nicotine dependence was assessed using the Fagerström Test for Nicotine Dependence (FTND) [23]. Scores on the FTND range from 0 to 10, with higher scores representing greater nicotine dependence. Participants also completed the Commitment to Quitting Smoking Scale (CQSS) [24] as a measure of their commitment to resist smoking in the face of difficulties, cravings, and negative affect; scores on the CQSS range from 1 to 5, with higher scores representing greater commitment to resist smoking. Smoking-specific experiential avoidance, or the tendency to smoke cigarettes in order avoid the experience of cigarette cravings, was measured with the physical sensations subscale [25] of the Avoidance and Inflexibility Scale (AIS) (adapted from Gifford et al [26]), which yields scores ranging from 1 to 5, with lower scores representing a greater avoidance of physical sensations that cue smoking. Participants were also asked to report on their use of e-cigarettes.

Symptoms of Depression and Anxiety Disorders

To assess for elevated levels depression and anxiety, participants completed validated self-report screening instruments for five common affective disorders. Symptoms of depression were assessed with the Center for Epidemiologic Studies Depression scale using the recommended score of ≥16 to indicate elevated symptoms [27]. Elevated symptoms of PTSD were indicated by the recommended score of ≥14 on the 6-item PTSD Checklist [28]. The Autonomic Nervous System Questionnaire was used to assess symptoms of PD; elevated PD symptoms were indicated if participants reported having ≥1 panic attacks within the past month with at least one occurring in a situation in which they were not in danger or the center of attention [29]. A score of ≥10 on the 7-Item Generalized Anxiety Disorder scale [30] indicated elevated symptoms of GAD. Finally, a score of ≥6 the Mini-Social Phobia Inventory [31] indicated elevated symptoms of SAD.

Smokers who screened positive for one or more affective disorders were classified into six nonmutually exclusive subgroups based on their baseline affective disorder symptomology: depression, PTSD, PD, GAD, SAD, and more than one affective disorder. Thus, these participants could be classified into more than one ADS subgroup. Participants who did not screen positive for any affective disorder listed above were categorized as not having elevated symptoms of any affective disorder (No AD).

Treatment Acceptability: Satisfaction and Use

As a subjective measure of treatment acceptability, participants answered three treatment satisfaction questions at the 3-month follow-up survey (eg, How useful did you find your assigned website?). Response options ranged from 1 (not at all) to 5 (very much), with a score ≥2 (ie, somewhat or more) indicating participant satisfaction. Additionally, website engagement over the 12-month period was calculated from data automatically logged by the secured server as indices of two objective measures of treatment acceptability—number of log-ins (primary use outcome) and total minutes spent on the website (secondary use outcome). We qualified user activity occurring more than 15 minutes after the last instance of activity as a new log-in.

Cessation Outcomes

Consistent with primary cessation outcomes reported in the parent trial, the primary cessation outcome for this study was self-reported 30-day point prevalence abstinence (PPA) (no smoking, not even a puff in the last 30 days) at 12 months postrandomization analyzed using a complete case methodology. The secondary cessation outcome also used 30-day PPA but with missing values imputed as smoking. Biochemical confirmation was not used in accordance with recommendations for assessing smoking status in large, population-based cessation trials in which there is no face-to-face contact and the demand characteristics for false reporting are minimal [32].

XSL•FO
RenderX

Statistical Analysis

For descriptive purposes, baseline characteristic variables of participants with ADS (each diagnostic group considered separately) were examined and compared against the no AD group using chi-square tests for categorical variables and *t* tests for continuous variables. Logistic regression models were used to evaluate differences in treatment satisfaction and cessation outcomes between each diagnostic group and the no AD group. Negative binomial models were used to evaluate differences in website use between groups. Models included variables used in stratified randomization for the main trial (ie, gender, education, and smoking more than 20 cigarettes per day), treatment arm, and treatment arm by diagnostic group interactions were considered. Nonsignificant interaction terms were dropped from final models. Analyses were completed using R version 3.4.2 (The R Foundation) and R package MASS [33], and all statistical tests were two-sided with alpha=.05.

Results

Participant Symptomology

Of the 2512 participants who provided sufficient data on the screening measures to establish whether the screen was positive or negative, the majority (1938/2512, 77.15%) screened positive for at least one affective disorder at baseline. Of those, the majority screened positive for depression (1470/1938, 75.85%), followed by PTSD (1383/1938, 71.36%), PD (1145/1938, 59.08%), GAD (903/1938, 46.59%), and SAD (797/1938, 41.25%). Additionally, as can be seen in Table 1, most of these participants screened positive for more than one affective disorder. Overall, 15.18% (355/2339) screened positive for one affective disorder, 13.59% (318/2339) screened positive for two, 16.76% (392/2339) screened positive for three, 15.01% (351/2339) screened positive for four, and 14.92% (349/2339) screened positive for all five affective disorders assessed in the study, suggesting high levels of elevated affective symptomology overall.

Baseline Characteristics

The baseline characteristics of all groups can be found in Table 2. Participants with ADS differed from those without ADS in a number of ways on demographic and smoking variables. Compared to smokers without ADS, all ADS subgroups were significantly younger, less educated, and more likely to identify as a sexual minority (ie, lesbian, gay, or bisexual). Some groups (GAD, SAD) comprised a greater proportion of women, and all ADS subgroups except the PD subgroup were less likely to be married. No differences were found between groups regarding race or ethnicity. Relative to the group without ADS, all ADS subgroups had significantly greater levels of nicotine dependence (all $P<.001$), avoidance of physical sensations that cue smoking (eg, urges) (all $P<.001$), and lower commitment to quitting scores (all $P<.001$).

Treatment Acceptability

As can be seen in Table 3, some differences emerged regarding subjective treatment acceptability ratings and objective website use data between those with and without ADS. Contrary to our predictions, relative to those who did not screen positive for any affective disorder, all ADS subgroups reported significantly greater satisfaction with their assigned website. Although differences in ratings of the other two indices of treatment acceptability (ie, usefulness of assigned website, whether or not they would recommend the website to a friend) were not significantly different, they were in the same direction such that those with ADS reported greater acceptability. No treatment by symptom group interaction was significant.

With regard to objective website use data, our hypotheses were partially supported. Relative to the no ADS group, the GAD group logged in to their assigned website significantly fewer times ($P<.001$). For all other symptom groups, there were significant treatment by symptom group interactions such that each ADS group logged in to the Smokefree.gov website significantly fewer times than the no ADS group. No significant differences in number of log-ins between groups were found in the ACT-based arm. All symptom subgroups spent significantly less time logged in to their assigned website compared with the no ADS group, and no treatment by symptom group interaction was significant.

Table 1. Frequency and proportion of participants screening positive for multiple affective disorders.

ADS subgroups	Depression (n=1470), n/N (%)	PTSD[a] (n=1383), n/N (%)	PD[b] (n=1145), n/N (%)	GAD[c] (n=903), n/N (%)	SAD[d] (n=797), n/N (%)
Depression	—	1152/1378 (83.59)	854/1143 (74.72)	836/900 (92.89)	697/795 (87.67)
PTSD	1152/1450 (79.45)	—	836/1144 (73.08)	796/900 (88.44)	664/796 (82.42)
PD	854/1343 (63.59)	836/1285 (65.05)	—	597/847 (70.48)	479/733 (65.35)
GAD	836/1364 (61.29)	796/1375 (57.89)	597/1139 (52.14)	—	506/796 (63.57)
SAD	697/1468 (47.48)	664/1382 (48.05)	479/1145 (41.83)	506/902 (56.10)	—

[a]PTSD: posttraumatic stress disorder.

[b]PD: panic disorder.

[c]GAD: generalized anxiety disorder.

[d]SAD: social anxiety disorder.

XSL•FO
RenderX

Table 2. Baseline characteristics by affective disorder symptomology group.

Characteristics	Comparison group (no ADS[a]; n=574)	Affective disorder symptomology subgroup					
		Depression (n=1470)	PTSD[b] (n=1383)	PD[c] (n=1145)	GAD[d] (n=903)	SAD[e] (n=797)	2+ ADS[f] (n=1529)
Demographics							
Age in years, mean (SD)	49.1 (13.0)	44.5 (13.5)[g]	43.5 (13.2)[g]	43.9 (13.1)[g]	42.3 (13.4)[g]	43.3 (13.4)[g]	43.9 (13.3)[g]
Female, n (%)	441 (76.8)	1181 (80.34)	1096 (79.25)	916 (80.00)	734 (81.28)[i]	653 (81.93)[i]	1226 (80.18)
Caucasian, n (%)	413 (71.9)	1051 (71.49)	962 (69.56)	813 (71.00)	633 (70.09)	559 (70.14)	1079 (70.57)
Hispanic, n (%)	43 (7.5)	134 (9.12)	137 (9.91)	107 (9.34)	96 (10.63)	80 (10.04)	148 (9.68)
Married, n (%)	232 (40.4)	494 (33.61)[h]	473 (34.20)[i]	423 (36.94)	315 (34.88)[i]	255 (31.99)[h]	533 (34.86)[i]
≤ HS[j] education, n (%)	121 (21.1)	454 (30.88)[g]	415 (30.00)[g]	330 (28.82)[g]	286 (31.67)[g]	266 (33.38)[g]	466 (30.48)[g]
LGB[k], n (%)	40 (6.9)	168 (11.43)[h]	159 (11.49)[h]	139 (12.14)[h]	115 (12.74)[g]	93 (11.67)[h]	175 (11.45)[h]
Smoking variables							
FTND[l], mean (SD)	5.4 (2.3)	5.8 (2.1)[g]	5.8 (2.2)[g]	5.8 (2.2)[g]	5.9 (2.1)[g]	5.9 (2.1)[g]	5.8 (2.2)[g]
>20 cpd[m], n (%)	174 (30.3)	520 (35.37)[i]	468 (33.84)	388 (33.89)	307 (33.99)	268 (33.62)	518 (33.88)
Past month e-cigarette use, n (%)	179 (31.2)	532 (36.19)[i]	506 (36.59)[i]	413 (36.07)	323 (35.77)	287 (36.01)	553 (36.17)[i]
CQSS[n], mean (SD)	4.2 (0.7)	3.9 (0.8)[g]	3.9 (0.8)[g]	4.0 (0.8)[g]	4.0 (0.8)[g]	3.9 (0.8)[g]	3.9 (0.8)[g]
AIS-sensations[o], mean (SD)	3.1 (0.5)	2.9 (0.5)[g]	2.8 (0.4)[g]	2.8 (0.4)[g]	2.8 (0.4)[g]	2.8 (0.4)[g]	2.8 (0.4)[g]

[a]ADS: affective disorders and symptoms.

[b]PTSD: posttraumatic stress disorder.

[c]PD: panic disorder.

[d]GAD: generalized anxiety disorder.

[e]SAD: social anxiety disorder.

[f]2+ ADS: screened positive for 2 or more affective disorders.

[g]*P*<.001.

[h]*P*<.01.

[i]*P*<.05.

[j]HS: high school.

[k]LGB: identify as lesbian, gay, or bisexual.

[l]FTND: Fagerström Test for Nicotine Dependence.

[m]cpd: cigarettes per day.

[n]CQSS: Commitment to Quitting Smoking Scale.

[o]AIS-sensations: Avoidance and Inflexibility Scale–sensations subscale.

XSL•FO
RenderX

Table 3. Treatment acceptability (subjective and objective indices) by affective disorders and symptoms group.

Outcome variable	Comparison group (no ADS[a])	Affective disorder symptomology subgroup					
		Depression	PTSD[b]	PD[b]	GAD[d]	SAD[e]	2+ ADS[f]
Subjective, n/N (%)							
Satisfied with website	359/473 (75.89)	938/1123 (83.53)[g]	910/1090 (83.48)[g]	734/891 (82.38)[h]	598/699 (85.55)[i]	540/640 (84.38)[g]	991/1184 (83.69)[g]
Website useful for quitting	334/483 (69.15)	835/1162 (71.86)	829/1126 (73.62)	664/918 (72.33)	527/714 (73.81)	487/654 (74.46)	892/1221 (73.05)
Would recommend website to friend	369/409 (90.22)	957/1021 (93.73)	920/986 (93.31)	738/793 (93.06)	604/638 (94.67)[h]	537/575 (93.39)	1008/1079 (93.42)
Objective							
Number of log-ins, mean (SD)	7.9 (19.9)	_[j]	–	–	5.8 (15.4)[i]	–	–
ACT[k] website							
Mean (SD)	9.3 (21.7)	8.6 (31.2)	8.6 (31.5)	8.4 (26.0)	—	9.6 (40.6)	8.5 (30.7)
n	270	733	690	561	—	389	759
Smokefree.gov							
Mean (SD)	6.7 (18.1)	4.6 (6.5)[i]	4.5 (6.2)[i]	4.5 (6.0)[i]	—[l]	4.5 (5.4)[i]	4.5 (6.3)[i]
n	304	737	693	584	—	408	770
Total minutes	35.4 (107.5)	29.2 (84.1)[i]	29.3 (86.4)[i]	26.9 (47.2)[i]	26.3 (70.3)[i]	31.4 (108.8)[i]	28.6 (82.7)[i]

[a]ADS: affective disorders and symptoms.

[b]PTSD: posttraumatic stress disorder.

[c]PD: panic disorder.

[d]GAD: generalized anxiety disorder.

[e]SAD: social anxiety disorder.

[f]2+ ADS: screened positive for 2 or more affective disorders.

[g]*P*<.01.

[h]*P*<.05.

[i]*P*<.001.

[j]Indicates significant treatment by subgroup interaction; results for these groups are separated by treatment arm in subsequent rows.

[k]ACT: acceptance and commitment therapy.

[l]Indicates no significant treatment by subgroup interaction.

Cessation Outcomes

Cessation rates for each ADS group are shown in Table 4. The 12-month complete-case quit rate among participants without ADS was descriptively higher than all other groups (153/520, 29.4%). The quit rate was statistically lower for some but not all ADS subgroups. Specifically, participants in the depression (306/1267, 24.15%; odds ratio [OR] 0.78, 95% CI 0.62-0.98; *P*=.03), PTSD (294/1215, 24.19%; OR 0.78, 95% CI 0.62-0.98; *P*=.03), PD (229/1003, 22.83%; OR 0.73, 95% CI 0.57-0.92; *P*=.009), and two or more ADS (323/1332, 24.25%; OR 0.78, 95% CI 0.62-0.98; *P*=.03) groups exhibited statistically lower quit rates of 20% to 25% relative magnitude than the no ADS group; those who screened positive for GAD (195/792, 24.6%; *P*=.08) and SAD (176/696, 25.3%; *P*=.13) did not. None of the treatment arm by symptom group interactions were significant. Models examining the secondary cessation outcome where missing values were imputed as smoking resulted in a similar pattern of results, with the exception that the difference for the GAD subgroup became significant.

XSL•FO
RenderX

Table 4. Logistic regressions comparing 12-month cessation outcomes by affective disorders and symptoms group.

Characteristic	30-day PPA[a], n/N (%)	Models controlling for trial stratification variables[b] and treatment arm	
		OR[c] (95% CI)	*P* value
Complete case analyses			
No ADS[d]	153/520 (29.42)	reference group	reference group
Depression	306/1267 (24.15)	0.78 (0.62-0.98)	.03
PTSD[e]	294/1215 (24.19)	0.78 (0.62-0.98)	.03
PD[f]	229/1003 (22.83)	0.73 (0.57-0.92)	.009
GAD[g]	195/792 (24.62)	0.80 (0.62-1.03)	.08
SAD[h]	176/696 (25.29)	0.82 (0.63-1.06)	.13
2+ ADS[i]	323/1332 (24.25)	0.78 (0.62-0.98)	.03
Missing=smoking analyses			
No ADS	153/574 (26.55)	reference group	reference group
Depression	306/1470 (20.82)	0.74 (0.59-0.93)	.009
PTSD	294/1383 (21.26)	0.76 (0.60-0.95)	.02
PD	229/1145 (20.00)	0.70 (0.56-0.89)	.004
GAD	195/903 (21.59)	0.77 (0.60-0.99)	.04
SAD	176/797 (22.08)	0.79 (0.61-1.01)	.06
2+ ADS	323/1529 (21.12)	0.75 (0.60-0.94)	.01

[a]PPA: point prevalence abstinence; no smoking, not even a puff in the last 30 days.

[b]Treatment stratification variables were gender, high school or less education, and smoking >20 cigarettes per day.

[c]OR: odds ratio.

[d]ADS: affective disorders and symptoms.

[e]PTSD: posttraumatic stress disorder.

[f]PD: panic disorder.

[g]GAD: generalized anxiety disorder.

[h]SAD: social anxiety disorder.

[i]2+ ADS: screened positive for 2 or more affective disorders.

Post Hoc Analyses

To develop a better understanding of the factors contributing to lower quit rates among the symptom subgroups who were less likely to quit, post hoc analyses examined (1) whether each ADS group uniquely predicts abstinence after including baseline variables that differed between groups and were associated with cessation in the models and (2) whether website use (ie, log-ins, total time spent on website) mediated the relationship between ADS subgroup and abstinence.

As shown in Table 5, after adding the additional baseline variables to logistic regression models for each ADS subgroup, only screening positive for PD remained significantly negatively associated with cessation after controlling for baseline levels of nicotine dependence and commitment to quitting (OR 0.78, 95% CI 0.61-0.99; *P*=.04) in the analyses in which missing values were imputed as smoking. All other differences became nonsignificant and may represent variables that are part of a causal pathway between affective disorders and difficulty quitting. Mediation models tested the hypothesis that website use might indirectly explain the relationship between depression, PTSD, PD, and 2+ ADS subgroups (versus the no ADS group) and the primary cessation outcome [34]. However, none of the estimated indirect effects were significant, suggesting that fewer log-ins and less time spent on the websites did not help explain the observed lower quit rates for these ADS subgroups in this study (data not shown).

XSL•FO
RenderX

Table 5. Post hoc logistic regression analyses including additional covariates.

Characteristic	Models controlling for trial potential confounding variables		
	OR[a] (95% CI)	*P* value	Additional covariates
Complete case analyses			
No ADS[b]	reference group	reference group	
Depression	0.90 (0.71-1.13)	.36	FTND[c], CQSS[d]
PTSD[e]	0.90 (0.71-1.14)	.39	FTND, CQSS
PD[f]	0.81 (0.63-1.03)	.09	FTND, CQSS
2+ ADS[g]	0.87 (0.69-1.10)	.26	FTND, CQSS
Missing=smoking analyses			
No ADS	reference group		
Depression	0.84 (0.67-1.06)	.15	FTND, CQSS
PTSD	0.87 (0.69-1.10)	.24	FTND, CQSS
PD	0.78 (0.61-0.99)	.04	FTND, CQSS
2+ ADS	0.86 (0.67-1.11)	.24	CQSS

[a]OR: odds ratio.

[b]ADS: affective disorders and symptoms.

[c]FTND: Fagerström Test for Nicotine Dependence.

[d]CQSS: Commitment to Quitting Smoking Scale.

[e]PTSD: posttraumatic stress disorder.

[f]PD: panic disorder.

[g]2+ ADS: screened positive for 2 or more affective disorders.

Discussion

Principal Findings

To inform future treatment development efforts for smokers with ADS, we conducted the first study comparing treatment acceptability and long-term smoking cessation outcomes between smokers who screened positive for five common affective disorders (depression, PTSD, PD, GAD, SAD, as well as those who screened positive for ≥2 diagnostic categories). All ADS subgroups differed from the no ADS group (those who did not screen positive for any affective disorder) on most demographic and smoking characteristics assessed at baseline, similar to other reports [19,35,36].

All ADS subgroups had higher subjective acceptability ratings of their assigned website compared with the no ADS group. It may be that, relative to smokers without affective disorder symptomology, smokers with ADS see greater potential for Web-based smoking interventions or find the Web-based format particularly helpful and convenient, leading to greater subjective acceptability among those groups. Although unanticipated, this finding is important and demonstrates that smokers who screen positive for affective disorders are open to this treatment modality and find them helpful. However, the higher subjective ratings did not translate into greater use of the websites as one might expect. Although most ADS subgroups logged in significantly less to the Smokefree.gov website (with no differential log-in rates for the ACT-based website) relative to the no ADS group, all ADS subgroups spent significantly less time on both websites overall, which is on par with our hypotheses. The differential log-in rates for the websites across groups may mean that smokers with ADS engage more with Web-based cessation programs (ie, find them more acceptable) when they are grounded in ACT, presented in a stepwise fashion, or when they have features that prompt tracking.

Reasons for the discrepancy in the findings between the subjective and objective indices of acceptability are unclear. On the one hand, it may be that the cognitive and affective symptoms experienced by smokers with ADS impede them from using the websites as much as individuals without ADS. Alternatively, given that the websites in the parent trial were designed for the general population of smokers, it may also be that smokers with ADS found them less personally relevant, which could have contributed to them having lower levels of engagement [37-40].

The quit rates at 12-month follow-up for all symptom subgroups were promising (20% to 25%). However, in the primary analyses most symptom subgroups (those screening positive for depression, PTSD, PD, or ≥2 affective disorders) had significantly lower quit rates (23% to 24%) relative to the no ADS group (29%). Thus, there is some evidence of the possibility of differential outcomes being disorder-specific. However, all subgroups had descriptively lower quit rates than the no ADS group, suggesting that there is room to improve quit rates for all symptom subgroups and doing so would likely be significant at the population level considering the large proportion of smokers who screen positive for ADS.

XSL•FO
RenderX

Results from post hoc analyses suggest that the lower cessation rates among most groups may be influenced in part by their having greater levels of nicotine dependence and lower commitment to resist smoking in difficult situations at baseline. While greater levels of nicotine dependence among individuals with ADS is a robust finding in the literature [6,19,41], this is the first study to our knowledge demonstrating that these individuals also have significantly lower levels of commitment to resist smoking during difficult situations, which may be an important target to include in smoking interventions designed for these groups. However, unlike in previous studies of Web-based smoking interventions where greater use was associated with better cessation outcomes [42-45], the lower rates of cessation in the affective disorder subgroups were not mediated by website use. It may be that higher engagement only mediates cessation when the intervention adequately addresses the needs of the users. If the intervention does not address the user's needs or is not perceived as relevant to the user, it follows that the user wouldn't be as engaged and that engagement, per se, wouldn't mediate cessation. Taken together, this suggests that increasing use alone among these groups may not be sufficient for improving cessation outcomes and that targeted interventions should also focus on addressing the unique, modifiable needs of smokers with affective conditions (eg, commitment to resist smoking in difficult situations) as a means of improving outcomes of these groups. Moreover, differences in nonmodifiable factors (eg, demographics) between these groups should also be considered when developing targeted interventions to increase their relevance for these groups. Future research is needed to determine if targeted Web-based interventions are associated with improved outcomes for these groups and if engagement mediates cessation outcomes in these cases.

Strengths and Limitations

There are numerous strengths to this study including a large, geographically diverse sample of adult smokers living in the United States, the use of validated self-report instruments to assess mental health conditions, and assessing long-term cessation outcomes at 12-months with a high rate of outcome data retention. Despite these strengths, the limitations of this study should be considered when interpreting the results. The primary limitation is the exploratory nature of the study as secondary analyses embedded in a larger clinical trial [21]; thus, our conclusions should be considered preliminary. A second limitation to the study is the use of self-report screening instruments to assess ADS. Individuals who screened positive for affective disorders may not all have the disorders they screened positive for and instead have elevated, subthreshold symptoms of the disorders. In line with previous research indicating that even individuals with subthreshold symptoms of affective disorders are at risk for continued smoking [3,46] and that screening instruments for affective disorders can be used to predict cessation [2], this study highlights that the screening instruments used in this study can be useful in identifying smokers who are less likely to quit even without a formal diagnostic assessment. Finally, we did not biochemically verify self-reported smoking abstinence at the 12-month follow-up. However, biochemical validation of abstinence is considered unnecessary for population-level intervention studies that otherwise have no face-to-face contact with participants because requiring this can bias study results [32].

Conclusions

This is the first study to compare treatment acceptability and long-term quit rates in response to Web-based smoking interventions between smokers who do versus do not screen positive for five common affective disorders. Findings suggest that while Web-based interventions are appealing to these groups, most individuals with ADS used their assigned website less and were less likely to quit smoking than their counterparts in response to the two interventions, each designed for the general population of smokers. Overall, these results support the need for developing targeted interventions for smokers with affective disorders and elevated affective symptomology.

Acknowledgments

The authors would like to acknowledge funding by the US National Cancer Institute (R01 CA166646; R01CA192849), and the US National Institute on Drug Abuse (R01 DA038411; F32 DA041100). Funders had no role in the trial conduct or interpretation of results.

Conflicts of Interest

In July 2016, JBB was a consultant to Glaxo Smith Kline, the makers of a nicotine replacement therapy. He now serves on the Scientific Advisory Board of Chrono Therapeutics, the makers of a nicotine replacement therapy device. JLH has received research support from Pfizer, the makers of a smoking cessation medication. Other authors have no declarations.

References

1. 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]
2. Cook BL, Wayne GF, Kafali EN, Liu Z, Shu C, Flores M. Trends in smoking among adults with mental illness and association between mental health treatment and smoking cessation. JAMA 2014 Jan 08;311(2):172-182 [FREE Full text] [doi: 10.1001/jama.2013.284985] [Medline: 24399556]

XSL•FO
RenderX

3. Landolt K, Ajdacic-Gross V, Angst J, Merikangas KR, Gamma A, Gutzwiller F, et al. Smoking and psychiatric disorders: have subthreshold disorders been overlooked? Nicotine Tob Res 2010 May;12(5):516-520. [doi: 10.1093/ntr/ntq023] [Medline: 20237120]
4. Lasser K, Boyd JW, Woolhandler S, Himmelstein DU, McCormick D, Bor DH. Smoking and mental illness: a population-based prevalence study. JAMA 2000;284(20):2606-2610. [Medline: 11086367]
5. West R, Evins AE, Benowitz NL, Russ C, McRae T, Lawrence D, et al. Factors associated with the efficacy of smoking cessation treatments and predictors of smoking abstinence in EAGLES. Addiction 2018 Aug;113(8):1507-1516 [FREE Full text] [doi: 10.1111/add.14208] [Medline: 29508470]
6. Ziedonis D, Hitsman B, Beckham JC, Zvolensky M, Adler LE, Audrain-McGovern J, et al. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health report. Nicotine Tob Res 2008 Dec;10(12):1691-1715. [doi: 10.1080/14622200802443569] [Medline: 19023823]
7. Callaghan RC, Veldhuizen S, Jeysingh T, Orlan C, Graham C, Kakouris G, et al. Patterns of tobacco-related mortality among individuals diagnosed with schizophrenia, bipolar disorder, or depression. J Psychiatr Res 2014 Jan;48(1):102-110. [doi: 10.1016/j.jpsychires.2013.09.014] [Medline: 24139811]
8. Prochaska JJ, Das S, Young-Wolff KC. Smoking, mental illness, and public health. Annu Rev Public Health 2017 Dec 20;38:165-185 [FREE Full text] [doi: 10.1146/annurev-publhealth-031816-044618] [Medline: 27992725]
9. Graham AL, Amato MS. Twelve million smokers look online for smoking cessation help annually: Health Information National Trends Survey Data, 2005-2017. Nicotine Tob Res 2019 Jan 04;21(2):249-252 [FREE Full text] [doi: 10.1093/ntr/nty043] [Medline: 29660037]
10. Graham AL, Carpenter KM, Cha S, Cole S, Jacobs MA, Raskob M, et al. Systematic review and meta-analysis of Internet interventions for smoking cessation among adults. Subst Abuse Rehabil 2016;7:55-69 [FREE Full text] [doi: 10.2147/SAR.S101660] [Medline: 27274333]
11. Haller CS, Etter J, Courvoisier DS. Trajectories in cigarette dependence as a function of anxiety: a multilevel analysis. Drug Alcohol Depend 2014 Jun 01;139:115-120. [doi: 10.1016/j.drugalcdep.2014.03.016] [Medline: 24703608]
12. Mañanes G, Vallejo MA. Usage and effectiveness of a fully automated, open-access, Spanish Web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2014;16(4):e111 [FREE Full text] [doi: 10.2196/jmir.3091] [Medline: 24760951]
13. Rabius V, Pike KJ, Wiatrek D, McAlister AL. Comparing internet assistance for smoking cessation: 13-month follow-up of a six-arm randomized controlled trial. J Med Internet Res 2008;10(5):e45 [FREE Full text] [doi: 10.2196/jmir.1008] [Medline: 19033146]
14. Seidman DF, Westmaas JL, Goldband S, Rabius V, Katkin ES, Pike KJ, et al. Randomized controlled trial of an interactive internet smoking cessation program with long-term follow-up. Ann Behav Med 2010 Feb;39(1):48-60. [doi: 10.1007/s12160-010-9167-7] [Medline: 20177844]
15. Graham AL, Papandonatos GD, Cobb CO, Cobb NK, Niaura RS, Abrams DB, et al. Internet and telephone treatment for smoking cessation: mediators and moderators of short-term abstinence. Nicotine Tob Res 2015 Mar;17(3):299-308 [FREE Full text] [doi: 10.1093/ntr/ntu144] [Medline: 25156528]
16. Muñoz RF, Barrera AZ, Delucchi K, Penilla C, Torres LD, Pérez-Stable EJ. International Spanish/English Internet smoking cessation trial yields 20% abstinence rates at 1 year. Nicotine Tob Res 2009 Sep;11(9):1025-1034 [FREE Full text] [doi: 10.1093/ntr/ntp090] [Medline: 19640833]
17. Kerkvliet JL, Wey H, Fahrenwald NL. Cessation among state quitline participants with a mental health condition. Nicotine Tob Res 2015 Jun;17(6):735-741. [doi: 10.1093/ntr/ntu239] [Medline: 25385874]
18. McFall M, Saxon AJ, Malte CA, Chow B, Bailey S, Baker DG, CSP 519 Study Team. Integrating tobacco cessation into mental health care for posttraumatic stress disorder: a randomized controlled trial. JAMA 2010 Dec 08;304(22):2485-2493 [FREE Full text] [doi: 10.1001/jama.2010.1769] [Medline: 21139110]
19. Piper ME, Cook JW, Schlam TR, Jorenby DE, Baker TB. Anxiety diagnoses in smokers seeking cessation treatment: relations with tobacco dependence, withdrawal, outcome and response to treatment. Addiction 2011 Feb;106(2):418-427 [FREE Full text] [doi: 10.1111/j.1360-0443.2010.03173.x] [Medline: 20973856]
20. Rogers ES, Smelson DA, Gillespie CC, Elbel B, Poole S, Hagedorn HJ, et al. Telephone smoking-cessation counseling for smokers in mental health clinics: a patient-randomized controlled trial. Am J Prev Med 2016 Apr;50(4):518-527. [doi: 10.1016/j.amepre.2015.10.004] [Medline: 26711163]
21. Bricker JB, Mull KE, McClure JB, Watson NL, Heffner JL. Improving quit rates of web-delivered interventions for smoking cessation: full-scale randomized trial of WebQuit.org versus Smokefree.gov. Addiction 2018 May;113(5):914-923. [doi: 10.1111/add.14127] [Medline: 29235186]
22. Watson NL, Mull KE, Heffner JL, McClure JB, Bricker JB. Participant recruitment and retention in remote ehealth intervention trials: methods and lessons learned from a large randomized controlled trial of two web-based smoking interventions. J Med Internet Res 2018 Aug 24;20(8):e10351 [FREE Full text] [doi: 10.2196/10351] [Medline: 30143479]
23. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]

XSL•FO
RenderX

24. Kahler CW, Lachance HR, Strong DR, Ramsey SE, Monti PM, Brown RA. The commitment to quitting smoking scale: initial validation in a smoking cessation trial for heavy social drinkers. Addict Behav 2007 Oct;32(10):2420-2424 [FREE Full text] [doi: 10.1016/j.addbeh.2007.04.002] [Medline: 17478057]
25. Bricker J, Wyszynski C, Comstock B, Heffner JL. Pilot randomized controlled trial of web-based acceptance and commitment therapy for smoking cessation. Nicotine Tob Res 2013 Oct;15(10):1756-1764 [FREE Full text] [doi: 10.1093/ntr/ntt056] [Medline: 23703730]
26. Gifford EV, Kohlenberg BS, Hayes SC, Antonuccio DO, Piasecki MM, Rasmussen-Hall ML, et al. Acceptance-based treatment for smoking cessation. Behav Ther 2004;35(4):689-705. [doi: 10.1016/S0005-7894(04)80015-7]
27. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
28. Lang AJ, Wilkins K, Roy-Byrne PP, Golinelli D, Chavira D, Sherbourne C, et al. Abbreviated PTSD Checklist (PCL) as a guide to clinical response. Gen Hosp Psychiatry 2012;34(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2012.02.003] [Medline: 22460001]
29. Stein MB, Roy-Byrne PP, McQuaid JR, Laffaye C, Russo J, McCahill ME, et al. Development of a brief diagnostic screen for panic disorder in primary care. Psychosom Med 1999;61(3):359-364. [Medline: 10367617]
30. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
31. Connor KM, Kobak KA, Churchill LE, Katzelnick D, Davidson JR. Mini-SPIN: a brief screening assessment for generalized social anxiety disorder. Depress Anxiety 2001;14(2):137-140. [Medline: 11668666]
32. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
33. Venables W, Ripley B. Modern Applied Statistics with S. Fourth Edition. New York: Springer; 2002.
34. Hayes AF. A primer on multilevel modeling. Human Comm Res 2006 Oct;32(4):385-410 [FREE Full text] [doi: 10.1111/j.1468-2958.2006.00281.x]
35. Anthenelli RM, Benowitz NL, West R, St Aubin L, McRae T, Lawrence D, et al. Neuropsychiatric safety and efficacy of varenicline, bupropion, and nicotine patch in smokers with and without psychiatric disorders (EAGLES): a double-blind, randomised, placebo-controlled clinical trial. Lancet 2016 Jun 18;387(10037):2507-2520. [doi: 10.1016/S0140-6736(16)30272-0] [Medline: 27116918]
36. Vickerman K, Schauer GL, Malarcher A, Zhang L, Mowery P, Nash C. Quitline use and outcomes among callers with and without mental health conditions: a 7-month follow-up evaluation in three states. Biomed Res Int 2015:1 [FREE Full text] [doi: 10.1155/2015/817298] [Medline: 26273647]
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. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2016 Dec 13;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
39. Tedeschi GJ, Cummins SE, Anderson CM, Anthenelli RM, Zhuang Y, Zhu S. Smokers with self-reported mental health conditions: a case for screening in the context of tobacco cessation services. PLoS One 2016 Jul;11(7):e0159127. [doi: 10.1371/journal.pone.0159127] [Medline: 27391334]
40. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2):81-92. [doi: 10.1097/NCN.0b013e3182065eed] [Medline: 21164337]
41. Trosclair A, Dube SR. Smoking among adults reporting lifetime depression, anxiety, anxiety with depression, and major depressive episode, United States, 2005-2006. Addict Behav 2010 May;35(5):438-443. [doi: 10.1016/j.addbeh.2009.12.011] [Medline: 20079577]
42. An LC, Schillo BA, Saul JE, Wendling AH, Klatt CM, Berg CJ, et al. Utilization of smoking cessation informational, interactive, and online community resources as predictors of abstinence: cohort study. J Med Internet Res 2008;10(5):e55 [FREE Full text] [doi: 10.2196/jmir.1018] [Medline: 19103587]
43. Bricker JB, Sridharan V, Zhu Y, Mull KE, Heffner JL, Watson NL, et al. Trajectories of 12-month usage patterns for two smoking cessation websites: exploring how users engage over time. J Med Internet Res 2018 Apr 20;20(4):e10143 [FREE Full text] [doi: 10.2196/10143] [Medline: 29678799]
44. Richardson A, Graham AL, Cobb N, Xiao H, Mushro A, Abrams D, et al. Engagement promotes abstinence in a web-based cessation intervention: cohort study. J Med Internet Res 2013;15(1):e14 [FREE Full text] [doi: 10.2196/jmir.2277] [Medline: 23353649]
45. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]

XSL•FO
RenderX

46. Watson NL, Heffner JL, McClure JB, Mull KE, Bricker JB. Differential prevalence of established risk factors for poor cessation outcomes among smokers by level of social anxiety. Am J Addict 2017 Mar;26(2):176-182 [FREE Full text] [doi: 10.1111/ajad.12509] [Medline: 28191916]

Abbreviations

ACT: acceptance and commitment therapy
AD: affective disorder
ADS: affective disorders and symptoms
AIS: Avoidance and Inflexibility Scale
CQSS: Commitment to Quitting Smoking Scale
FTND: Fagerström Test for Nicotine Dependence
GAD: generalized anxiety disorder
LGB: lesbian, gay, or bisexual
PD: panic disorder
PPA: point prevalence abstinence
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial
SAD: social anxiety disorder

Edited by G Eysenbach; submitted 25.01.19; peer-reviewed by M Amato, Y Zhu, T Tanigawa; comments to author 14.03.19; revised version received 02.04.19; accepted 20.04.19; published 19.06.19.

Please cite as:
Watson NL, Heffner JL, Mull KE, McClure JB, Bricker JB
Comparing Treatment Acceptability and 12-Month Cessation Rates in Response to Web-Based Smoking Interventions Among Smokers Who Do and Do Not Screen Positive for Affective Disorders: Secondary Analysis
J Med Internet Res 2019;21(6):e13500
URL: https://www.jmir.org/2019/6/e13500/
doi:10.2196/13500
PMID:31219052

©Noreen L Watson, Jaimee L Heffner, Kristin E Mull, Jennifer B McClure, Jonathan B Bricker. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobility-Focused Knowledge Translation Randomized Controlled Trial to Improve Physical Activity: Process Evaluation of the Move4Age Study

Sarah E Neil-Sztramko[1], PhD; Jenna Smith-Turchyn[2], PhD; Julie Richardson[3], PhD; Maureen Dobbins[4], BScN, PhD

[1]School of Nursing, McMaster University, Hamilton, ON, Canada

[2]Faculty of Kinesiology & Physical Education, University of Toronto, Toronto, ON, Canada

[3]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada

[4]National Collaborating Centre for Methods and Tools, School of Nursing, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Sarah E Neil-Sztramko, PhD
School of Nursing
McMaster University
175 Longwood Road South, Suite 210a
Hamilton, ON, L8P0A1
Canada
Phone: 1 9055259140 ext 20459
Email: neilszts@mcmaster.ca

Abstract

Background: Maintaining physical activity and physical function is important for healthy aging. We recently completed a randomized controlled trial of a targeted knowledge translation (KT) intervention delivered through the McMaster Optimal Aging Portal with the goal to increase physical activity and physical mobility in middle-aged and older adults, with results reported elsewhere.

Objective: The purpose of this process evaluation study is to explore which KT strategies were used by both intervention and control group participants, as well as the intervention groups' engagement, satisfaction, and perceived usefulness of the targeted KT intervention.

Methods: Data on engagement with the intervention materials were gathered quantitatively through Google Analytics and Hootsuite throughout the intervention. Qualitative data were collected through a combination of open-ended surveys and qualitative interviews with a subset of participants at the end of the study to further understand engagement, satisfaction, and usefulness of the KT strategies.

Results: Throughout the intervention period, engagement with content delivered through weekly emails was highest, and participants rated email content most favorably in both surveys and interviews. Participants were generally satisfied with the intervention, noting the ease of participating and the distillation of information in an easy-to-access format being beneficial features. Participants who did not find the intervention useful were those with already high levels of baseline physical activity or physical function and those who were looking for more specific or individualized content.

Conclusions: This process evaluation provides insight into our randomized controlled trial findings and provides information that can be used to improve future online KT interventions.

Trial Registration: ClinicalTrials.gov NCT02947230; https://clinicaltrials.gov/ct2/show/nct02947230 (Archived by WebCite at http://www.webcitation.org/78t4tR8tM)

(J Med Internet Res 2019;21(6):e13965) doi:10.2196/13965

KEYWORDS

process evaluation; physical activity; health information; mobility; older adults; knowledge translation

XSL•FO
RenderX

Introduction

Maintaining physical mobility is important for healthy aging and maintaining functional independence [1]. Regular physical activity has been shown to slow the age-related decline in physical mobility [2] and is associated with decreased risk of mobility disability and mortality [1,3,4]. Participation in regular physical activity has been shown to have a number of positive effects on functional, metabolic, cardiovascular, and cognitive outcomes, as well as improvements in quality of life [2]. Given these beneficial effects of exercise, effective knowledge translation (KT) interventions that disseminate information to middle-aged and older adults on strategies that promote physical mobility are warranted.

Population-based survey data from Canada suggest that older adults do use the Internet to seek out health information [5]. One potential benefit of health resources disseminated online is the potential for a wide reach to a diverse audience. However, it may be difficult for members of the public to adequately identify trustworthy online resources [6]. In response to this problem, the McMaster Optimal Aging Portal (the Portal) was created in 2014 to serve as a publicly available repository of high-quality, evidence-based information about healthy aging [7,8]. User characteristics [9] and usability [10,11] of the Portal have been previously reported. While it is clear that people are using the Portal, questions remain as to whether use of the Portal has any impact on knowledge, health behaviors, or health outcomes.

Recent systematic reviews suggest that electronic behavior change interventions may have positive effects on physical activity and other health outcomes [12,13]. A review of information technology-based interventions on patient engagement and behavior change found that while IT platform-based interventions had positive effects on health outcomes, there are few published reports of process data on engagement and usability within the interventions [12].

Our team recently completed a randomized controlled trial of a targeted KT intervention through the Portal, aimed at improving physical activity and physical mobility in middle-aged and older adults (to be published elsewhere, currently under review). In line with previously published recommendations on process evaluations alongside randomized controlled trials [14], the purpose of this process evaluation was to explore participants' engagement, satisfaction, and perceived usefulness of the tailored KT intervention and the Portal in general. This provided a more in-depth understanding of our quantitative trial findings.

Methods

Trial Description

A full description of methods and results for this randomized controlled trial will be reported elsewhere. The Hamilton Integrated Research Ethics Board approved the study protocol, and all participants provided written informed consent. In brief, 510 participants—primarily female (430/510, 84.3%); mean age 64.7 years (SD 8.3; see Table 1)—were recruited online and randomized to a targeted KT intervention or self-serve control group. Approximately one-third were regular users of the Portal before the study, while one-third had never heard of the Portal before. Questionnaires were completed using an online survey platform at baseline (ie, prior to randomization), at the end of the 12-week study, and at 3-months postintervention. While there were no differences at the end-of-study or follow-up time points between groups for physical activity or self-monitoring, the intervention group did report more positive intentions to engage in mobility-related health behaviors and more favorable attitudes toward physical activity than the control group. In planned subgroup analyses to explore the effect of the intervention by participant characteristics, there was a significant intervention effect found among participants with low self-rated health.

Intervention Description

During the intervention period, those in the intervention group received a targeted intervention consisting of the following:

1. Mobility-focused weekly emails, which included links to blog posts (ie, short articles providing recent scientific evidence in a narrative format), evidence summaries (ie, 1-2-page documents describing findings from a systematic review in lay language), and Web-resource ratings (ie, evaluations to assess quality of existing third-party websites) on a weekly topic related to physical activity and/or mobility.
2. Social media posts via Twitter and Facebook, using the study-specific hashtag #Move4Age to highlight relevant information related to physical activity or mobility. Participants were initially invited to follow social media feeds at the beginning of the intervention and were reminded throughout the intervention period via email.
3. Invitation to visit the *Mobility and Physical Functioning* page on the Portal.

As the Portal is a publicly available website, control group participants were able to access the Portal during the intervention period, including the Portal's general weekly email alert subscription service; thus, our study did not include a *true* control group. The control group participants did not receive the targeted KT strategies, as described above.

The targeted KT intervention was informed by the theory of planned behavior [15]. The theory of planned behavior suggests that intention to engage in a particular behavior is an immediate precursor of the behavior, and that intention is based on attitude toward the behavior, subjective norms, and perceived behavior control. Through this intervention, we aimed to modify individuals' attitudinal beliefs through the provision of high-quality, evidence-based information about increasing physical activity and improving mobility and physical function. The educational materials provided were targeted at middle-aged and older adults and included actionable messages within the content, specifically within the blog posts [16], to act on normative and control beliefs (see Figure 1).

XSL•FO
RenderX

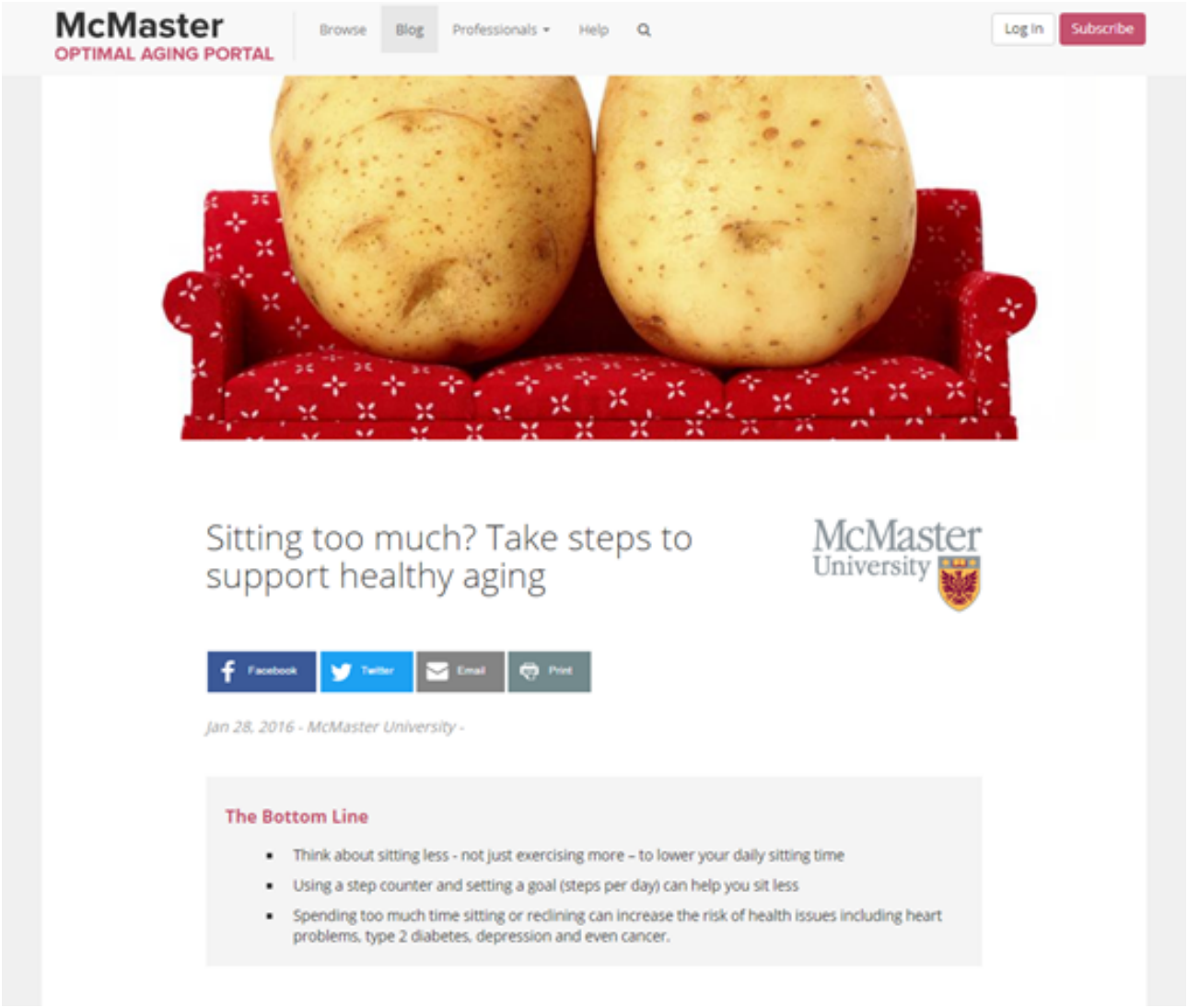
Table 1. Participant characteristics.

Characteristics	Total (N=510)	Intervention (n=256)	Control (n=254)	*P* value
Age (years), mean (SD)	64.7 (8.3)	64.7 (8.5)	64.6 (8.2)	.94
Gender, n (%)				.69
Male	80 (15.7)	38 (14.8)	42 (16.5)	
Female	430 (84.3)	218 (85.2)	212 (83.5)	
Education, n (%)				.91
High school diploma or less	36 (7.1)	18 (7.0)	18 (7.1)	
College diploma	111 (22.0)	58 (23.1)	53 (20.9)	
Bachelor's degree	217 (43.1)	104 (41.4)	113 (44.7)	
Postgraduate degree	140 (27.8)	71 (28.3)	69 (27.3)	
Employment status, n (%)				.19
Retired	304 (59.7)	157 (61.6)	147 (57.9)	
Full-time employment	121 (23.8)	60 (23.5)	61 (24.0)	
Part-time employment	65 (12.8)	28 (11.0)	37 (14.6)	
Long-term disability	6 (1.2)	1 (0.4)	5 (2.0)	
Other	13 (2.6)	9 (3.5)	4 (1.6)	
Geography, n (%)				.55
Urban	422 (82.7)	209 (81.6)	213 (83.9)	
Rural	74 (14.5)	41 (16.0)	33 (13.0)	
Not reported	14 (2.7)	6 (2.3)	8 (3.1)	
Self-rated health: *Excellent* or *Very Good*, n (%)	303 (59.4)	144 (56.3)	159 (62.6)	.07
Chronic disease, n (%)	283 (55.7)	141 (55.3)	142 (56.1)	.92
Falls				
Had a fall in the last 6 months, n (%)	103 (20.2)	41 (16.0)	62 (24.4)	.02
Number of falls, mean (SD)	1.6 (1.2)	1.4 (0.9)	1.7 (1.3)	.19
Visited a health care provider because of a fall, n (%)	35 (33.3)	15 (36.6)	20 (31.2)	.72
Previous Portal[a] use, n (%)				.98
Never used	172 (33.8)	87 (34.0)	85 (33.6)	
Regular user	153 (30.1)	76 (29.7)	77 (30.4)	
Used occasionally	184 (36.1)	93 (36.3)	91 (36.0)	
Sought information about improving mobility from a health care provider or other source in the last year, n (%)	220 (43.1)	118 (46.1)	102 (40.2)	.21

[a]Portal: McMaster Optimal Aging Portal.

XSL•FO
RenderX

Figure 1. Example of intervention material delivered through the McMaster Optimal Aging Portal blog.



Data Collection

Data on engagement with intervention materials delivered through the mobility-focused weekly emails were collected from participants in the intervention group only. A custom campaign was created for each week of the intervention using Google Analytics. This provided data for each weekly email on the number of participants who opened a link within the email, number of website sessions, number of page views, number of pages viewed per session, time per session, and bounce rate (ie, the proportion of individuals who only viewed one page per session).

Intervention materials disseminated through social media (ie, Twitter and Facebook) were identified using a study-specific hashtag, #Move4Age, and thus were publicly available. Number of shares, likes, and links clicked (Facebook) as well as number of retweets, likes, and URL clicks (Twitter) were collected using Hootsuite (Hootsuite Inc).

Data on participant satisfaction with, and perceived usefulness of, the KT strategies for the intervention group only, and the Portal itself for both groups, were collected at the end-of-study and 3-month follow-up time points using a combination of Likert scales ranging from 1 (strongly disagree) to 7 (strongly agree) and open-ended questionnaires. A subgroup of 50 participants also consented to take part in qualitative interviews following the end-of-study data collection. Semistructured interviews were conducted by a trained interviewer who was not involved in any other aspect of the study. Interviews were recorded and transcribed verbatim.

Data Analysis

Data on participant engagement with intervention materials is presented descriptively as mean and standard deviation as well as frequency and percentage where appropriate. Perceived satisfaction and usefulness of the Portal was compared between the intervention and control groups using an independent-samples *t* test for continuous variables and a chi-square test for categorical variables using SAS 9.4 (SAS Institute Inc).

Qualitative data from interview transcripts was entered into NVivo 11 (QSR International) for storage, indexing, searching, and coding. Two researchers (SENS and JST) reviewed a subset of interviews in duplicate to reach consensus on a coding scheme. Once agreement was reached, a thematic analysis was undertaken by the two researchers independently. Emergent

XSL•FO
RenderX

themes were compared to open-ended survey questions and quantitative study results to provide a deeper understanding of our quantitative study findings.

Results

Engagement and Satisfaction With Knowledge Translation Strategies

During the intervention period, 94.7% of intervention participants (198/209) reported receiving mobility-focused weekly emails. Engagement with email content was highest at the beginning of the study and declined throughout the course of the 12-week intervention (see Figure 2). Due to a technical issue, data on the number of emails successfully delivered and opened was not available. On average, one-third of intervention participants clicked through to the Portal from an email each week, ranging from a low of 17.6% in week 8 to 51.2% in week 2. Overall, engagement was highest in week 1 (766 total page views with an average of 4.07 pages and 5 minutes 32 seconds per session) and lowest in week 10 (117 total page views, 1.71 pages per session, and 1 minute 30 seconds per session). An increase in engagement was seen in the last week of the study: 218 total page views, 2.6 pages per session, and 2 minutes 36 seconds per session (see Table 2).

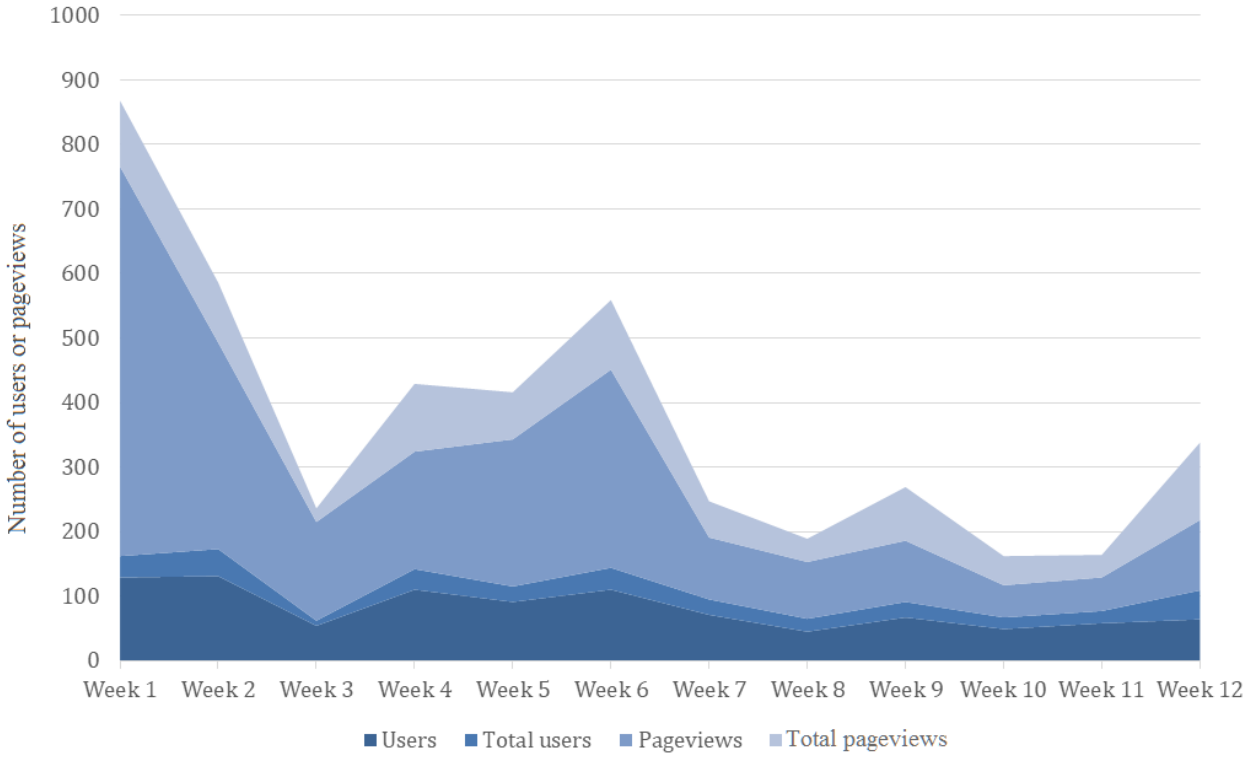
In qualitative interviews, participants reported that emails were the primary source of information utilized during the intervention period. With respect to positive aspects of the targeted KT intervention, a common theme emerged related to the ease of access to the study information. Participants reported that automatically receiving the mobility-focused content in their email inbox and having the large amount of information distilled in an easy-to-read format was appreciated.

> *I liked that it was information that came to me proactively. So I didn't have to always go looking for it. It was also in bite-sized chunks. It was information that came in like, in sort of manageable pieces of time and information, and they offered you skills that you could develop pretty easily and quickly, it wasn't a whole program you had to undertake, and it wasn't, it was sort of easy pieces to fit into my life.*

A related theme emerged around ease of selecting relevant content. Participants discussed selecting particularly relevant topics to read through, rather than reading through all information sent, which was viewed as a positive aspect of the intervention.

> *[My mother] is 84, I am 53, so I mean I look at my e-mail quite often, so, you know, I could see it, I knew it was there...sometimes I liked looking at the headings of what the evidence was, and you know I would peruse through and, you know, sometimes it was interesting to me, sometimes it was not—you know, but that was fine. I just deleted it, it was easy to delete if I wasn't that interested in it.*

Figure 2. Intervention participant engagement with email content during the 12-week study period.



XSL•FO
RenderX

Table 2. Intervention group engagement with mobility-focused email alerts during the 12-week intervention period.

Week	Topic	Unique users (N=256), n (%)	Total sessions, n	Total page views, n	Pages per session, n	Time per session, minutes, seconds	Bounce rate, %
1	Introduction to Move4Age	129 (50.4)	188	766	4.1	5, 32	35.1
2	Walking	131 (51.2)	209	493	2.5	3, 42	42.4
3	Enhancing social support with walking groups	54 (21.1)	78	215	2.5	2, 31	49.1
4	Balance	110 (43.0)	136	324	2.4	3, 8	42.9
5	Strength training	91 (35.5)	135	343	2.5	2, 55	43.4
6	Falls and injury prevention	110 (43.0)	152	451	2.9	3, 0	44.9
7	Maintaining a healthy body weight	71 (27.7)	95	191	2.2	3, 13	50.4
8	Using technology for self-monitoring	45 (17.6)	69	153	2.4	2, 13	51.4
9	Reducing sedentary time	67 (26.2)	86	186	1.0	1, 24	56.0
10	Alternative forms of exercise for mobility	49 (19.1)	63	117	1.7	1, 30	60.0
11	Cognition and mobility	58 (22.7)	81	129	1.7	1, 33	67.6
12	Overcoming physical limitations	64 (25.0)	85	218	2.5	2, 36	37.7
All weeks, mean (SD)		83.2 (30.8)	114.8 (46.4)	298.8 (183.2)	2.4 (0.7)	2, 46 (1, 5)	48.4 (9.0)

Only 7.7% (16/209) of intervention group participants reported using Twitter and 19.6% (41/209) reported using Facebook. Data on social media engagement by week is displayed in Table 3. During the study period, there were a total of 50 tweets marked with #Move4Age, ranging from 1 to 7 per week. These tweets garnered a total of 25,006 impressions (187-1485 impressions per tweet), 100 retweets (0-10 per tweet), 96 likes (0-10 per tweet), and 217 URL clicks (0-17 per tweet). The highest level of engagement was during weeks 3 and 4. The total number of study-specific Facebook posts was 15 (0-2 per week). These posts garnered 23,635 unique post impressions (278-3532 per post), 158 shares (1-28 per post), 298 likes (3-48 per post), 16 comments (1-4 per post), and 1206 link clicks (6-369 per post). The highest level of engagement was during weeks 1 and 6.

Table 3. Social media engagement throughout the 12-week study period.

Week	Twitter engagement					Facebook engagement				
	Tweets, n	Impressions per tweet, mean	Retweets per tweet, mean	Likes per tweet, mean	URL clicks per tweet, mean	Posts, n	Impressions per post, mean	Shares per post, mean	Likes per post, mean	Comments per post, mean
1	2	726.0	2.0	2.5	7.5	2	2359.5	15.5	32.5	1.0
2	7	260.9	0.3	0.9	3.3	1	2457.0	16.0	32.0	2.0
3	5	426.4	2.0	1.4	5.0	2	1042.5	6.5	14.0	1.5
4	4	811.8	4.3	4.3	10.8	2	1174.0	2.5	11.5	1.0
5	3	505.3	3.0	2.3	3.7	1	278.0	2.0	3.0	0.0
6	6	608.0	2.2	1.7	2.7	2	3439.0	28.0	38.0	4.0
7	6	420.3	1.7	1.3	4.0	1	513.0	3.0	7.0	0.0
8	7	397.0	1.9	1.7	2.9	1	1596.0	13.5	24.5	2.0
9	2	718.0	2.5	3.0	7.0	1	604.0	5.0	13.0	1.0
10	4	695.3	3.5	2.8	3.5	0	N/A[a]	N/A	N/A	N/A
11	3	275.3	0.0	2.0	1.3	1	1845.0	10.0	13.0	1.0
12	1	841.0	3.0	1.0	8.0	1	2155.0	18.0	27.0	0.0
All weeks, mean (SD)	4.2 (2.0)	557.1 (202.9)	2.2 (1.2)	2.1 (1.0)	5.0 (2.8)	1.3 (0.6)	1587.5 (1035.0)	10.9 (8.4)	19.6 (12.5)	1.2 (1.2)

[a]N/A: not applicable; there were no Facebook posts in week 10.

XSL•FO
RenderX

Control Group Engagement With the Portal

During the intervention period, 89.4% of control group participants reported registering for the Portal's regular weekly email alert subscription service, 4.6% reported following the Portal on Twitter, 19.4% reported following the Portal on Facebook, and 29.6% reported browsing for mobility-related content on the Portal (see Table 4). There were no significant differences in engagement with the different strategies between intervention and control group participants. Fewer control group participants reported that information from the Portal influenced a decision they made about physical activity (54.5% vs 68.0%, *P*=.006). Of those that reported that the Portal did influence a decision, the control group reported that this occurred less often (mean 2.73, SD 1.90 vs mean 3.43, SD 2.06, where 1 is *not often* and 7 is *very often*, *P*<.001).

Perceived Usefulness of the Intervention

At the end of study, participants reported mobility-focused emails to be useful (mean 5.27, SD 1.52, on a 1-7-score Likert scale) and reported favorably regarding their likelihood of continuing to subscribe (mean 5.46, SD 1.78) and recommending to a friend or family member (mean 5.29, SD 1.81; see Table 4). Of those who did use social media, responses were similarly favorable for usefulness of both Twitter and Facebook, likelihood of continued use, and likelihood of recommending to family or friends. The overall satisfaction with the intervention itself and with particular KT strategies was echoed by intervention group participants in qualitative interviews; however, a number of divergent themes emerged among participants with respect to the perceived usefulness and potential impact of the intervention. A first group of participants reported that they learned something new during the intervention that resulted in them making lifestyle changes around physical activity.

Table 4. Participant satisfaction and Portal[a] use during the 12-week intervention period.

Portal activity and influence	Intervention (n=209)	Control (n=216)	*P* value
Weekly email alerts from the Portal			.06
Received weekly email alerts, n (%)	198 (94.7)	193 (89.4)	
Mobility-specific email alerts are a useful strategy, mean (SD)	5.27 (1.52)[b]	N/A[c]	
Would continue to subscribe, mean (SD)	5.46 (1.78)	N/A	
Would recommend to a friend or family member, mean (SD)	5.29 (1.81)	N/A	
Portal access via Twitter			.27
Accessed the Portal via Twitter, n (%)	16 (7.7)	10 (4.6)	
Twitter is a useful strategy, mean (SD)	5.07 (1.87)	N/A	
Will continue to use, mean (SD)	6.12 (1.41)	N/A	
Would recommend to a friend or family member, mean (SD)	5.56 (1.50)	N/A	
Portal access via Facebook			.99
Accessed the Portal via Facebook, n (%)	41 (19.6)	42 (19.4)	
Facebook is a useful strategy, mean (SD)	5.61 (1.43)	N/A	
Will continue to use, mean (SD)	5.90 (1.30)	N/A	
Would recommend to a friend or family member, mean (SD)	5.32 (1.65)	N/A	
***Mobility and Physical Functioning* browse page**			.34
Used the *Mobility and Physical Functioning* browse page, n (%)	72 (34.4)	64 (29.6)	
Mobility-specific browse page is a useful strategy, mean (SD)	5.60 (1.10)	N/A	
Will continue to use, mean (SD)	5.51 (1.35)	N/A	
Would recommend to a friend or family member, mean (SD)	5.30 (1.60)	N/A	
Portal information influenced a decision about physical activity			
Number of participants who answered *yes*, n/N (%)[d]	140/206 (68.0)	116/213 (54.5)	.006
How often? mean (SD)	3.43 (2.06)	2.73 (1.90)	<.001

[a]Portal: McMaster Optimal Aging Portal.

[b]Numerical questions were answered on a scale of 1 (not often) to 7 (very often).

[c]N/A: not applicable.

[d]There were missing data (n=6) from this question: intervention (n=3) and control (n=3).

XSL•FO
RenderX

> *Certainly, I have been upping my exercises that involve pounding but stress on the muscles and the bone- the reason swimming doesn't do it because you have no impact. I certainly have continued to focus on impact activity.*
>
> *I was not previously aware that I needed to walk faster than a pleasant stroll. Now I am aware and each night after dinner my husband and I walk a fast 25 minutes.*

A second group of participants described the intervention materials as serving an important reminder to engage in health behaviors or reinforcement, but the content did not contain a lot of new information or result in new knowledge gained.

> *Yeah, in the sense that it just alerted me and kept me, kept me, ah, on target with my workouts and my walks, bicycling, and all that kind of stuff.*
>
> *For example, bone density is an issue for myself and it really wasn't always new information. It really more confirmed what I always researched and found out. Whether I've used it or not...um, for example, you're not supposed to swim, for example, as it's not an activity that increases bone density and I think I read that somewhere, but I already knew that. I'm not a totally uninformed consumer.*

A third group of participants noted that while they were satisfied with the intervention materials and content, they found the intervention not particularly impactful because they were already active or had no mobility limitations.

> *Yeah, not for any reason, ah, I'm relatively mobile, relatively mobile myself, but, um, that becomes an issue as you age and I suppose it's better to know about it before it's an issue, so I found it very interesting.*
>
> *You know, that again, I was pretty mobile, I had no, no issues beforehand, and I still don't have any issues, so although it didn't improve, it's because it would have been pretty hard for it to improve, I think.*

A final group of participants reported that the intervention itself had no impact and that they were generally disappointed with the intervention. Two prominent subthemes emerged within this group. Firstly, participants were dissatisfied with the intervention because the information provided was not specific enough.

> *It didn't really...it was superficial. It didn't tell you what to do, where to go...it was kind of information that's out there everywhere. There was nothing really new. I read it a couple of times and thought I'm missing something here...and then at some point I stopped reading them because I thought I would just glance over it because I thought that, I want the meat, okay? I don't want any more of these studies and this here and that...nothing gets in my pocket. My pocket meaning...I'm not getting any services.*

Secondly, some participants reported that the intervention seemed more appropriate for individuals with lower levels of baseline health or fitness.

> *I had a sense that it was targeted at people with already quite limited mobility, and not including those who had perhaps re-achieved a higher level of mobility through their own initiative.*
>
> *A lot of it seemed to be directed at people that had much more significant problems than me, so I'm pretty active and quite healthy, and so I was looking for things that would sort of help me stay that way or any tips, or any new information. And, yeah, I felt that a lot of the information, not all of it, but a lot of it was directed at people that already had significant problems.*

Discussion

Findings from our randomized controlled trial suggest that a targeted KT intervention may have a positive impact on levels of physical activity in middle-aged and older adults, particularly those with low baseline self-rated health. Findings from the process evaluation presented here help to understand these findings and provide guidance on the design and delivery of future KT interventions, particularly those using an online platform such as ours. To our knowledge, this is the first process evaluation conducted of an online KT intervention targeted at middle-aged and older adults.

Despite the large number of intervention participants who reported receiving the mobility-focused weekly emails, actual engagement with the Portal content as measured through Google Analytics was much lower. The proportion of participants in the intervention group who visited the Portal through a link within an email ranged from 17.6% to 51.2% (weeks 8 and 2, respectively). While lower than our target, these are still much higher than industry averages for health services email campaigns, which report average email open rates of 19.2% and click-through rates of only 6.4% [17]. In addition, those who did click through seem to be well engaged, which is reflected in the average number of pages viewed per session and the average time per session. The average length per session of 2 minutes and 46 seconds suggests that participants may have read the majority of the article on the page they visited. This hypothesis was confirmed in our qualitative interviews, where participants reported only clicking on links they were interested in reading and reading through the content on pages they visited. To our knowledge, few other similar studies have been conducted to provide comparative data. In a recent process evaluation of usage data from a publicly available, Web-based mental health portal for youth, 65.4% of sessions were less than one minute in duration [18]. In contrast, a Web-based intervention of interactive self-monitoring modules for women to prevent weight gain reported a median session duration of 12 minutes and 54 seconds [19]. However, it is difficult to directly compare session length, due to potential differences in the amount of content, education levels, and background knowledge of participants.

Qualitative data provided some explanation as to why such a low percentage of participants clicked through from each weekly email, despite rating the emails as highly useful. Participants reported that one of the benefits of receiving the weekly email

XSL•FO
RenderX

alerts was that it allowed them to quickly and easily identify personally relevant and topical material to read more thoroughly. We hypothesize that the length per session is reflective of participants being more interested and engaged with the content to which they chose to click through from the original email, while individuals did not click through to content that they were less interested in. In a process evaluation of a sexual health website, average time on the site ranged from 2 minutes and 7 seconds to 6 minutes and 36 seconds, depending on keywords searched and the referring website, demonstrating the substantial variability in usage that can occur within the same site depending on the topic at hand [20]. The intervention period did not occur over any major holidays, and we cannot identify any external reasons that would contribute to increased uptake in particular weeks (eg, week 6). To minimize participant burden, we did not collect feedback on weekly intervention content, although this information would be useful in further understanding what drives weekly variation in intervention engagement.

Given that there were no significant differences found in the proportion of participants in both the intervention and control groups that reported receiving weekly email alerts, following the Portal on social media, and following the mobility landing page, it is not surprising that no significant between-group differences were found at the end-of-study or postintervention follow-up data collection. Our findings that both groups self-reported significant increases in physical activity and self-monitoring mobility over time may indicate that the Portal's currently available KT strategies are sufficient to elicit behavior change, at least in some individuals. However, given the variation in change in physical activity across participants, and the still low proportion of participants meeting physical activity guidelines at the end of the study, it may also be that more specific tailoring of intervention materials is needed to see greater behavior change in some individuals.

As discussed above, a major theme that emerged from qualitative interviews was the benefit of ease of access to content through the weekly email alerts, and the *filtering* of relevant information. It may be that control group participants who signed up for this study due to their interest in physical activity and mobility also filtered the general Portal content in the same way. We were not able to track usage data in the form of number of clicks in email alerts through the control group, so we are not able to evaluate whether there was a difference in engagement between the targeted intervention email alerts and the general Portal email alerts to confirm this hypothesis.

Conversely, we also identified a group of participants within the qualitative interviews who were dissatisfied with the intervention due to its focus on general information and lack of specific instruction or resources. These findings are similar to those from our previous cross-sectional survey of Portal users, where a thematic analysis of open-ended questions identified limitations of the Portal being that information was not specific or in-depth enough [9]. This is an ongoing challenge for those designing online resources and delivering online KT interventions. These online resources must balance the logistics and feasibility of disseminating a broad-reaching behavior change intervention, while also being relevant to an individual. In a recent systematic review of eHealth interventions for physical activity promotion in older adults, four of the six website-based studies found a significant increase in physical activity compared to a no-intervention control group [13]. Each of these studies used some type of interactive component with intervention participants, such as a personal coach [21], an interactive website feature to track physical activity [22,23], or provision of a pedometer or accelerometer for monitoring physical activity [23,24]. While found to be effective, these components may not be feasible on a large-scale, publicly available resource such as the Portal. In a systematic review of strategies to facilitate use of Internet-delivered, health behavior change interventions in younger adults, Crutzen et al found that tailored communications, reminders, and incentives are useful strategies to increase user engagement with intervention content [25]. Qualitative data from our participants highlighting the greater engagement with personally relevant content support the use of tailoring in delivering intervention content. Future work may explore whether more specific tailoring of the intervention materials provided, such as by baseline values, knowledge, or preferences, may elicit greater overall behavior change and satisfaction with the KT intervention. Perhaps a theory-driven approach to tailoring, for example, using the Transtheoretical Model, could be explored to tailor intervention material based on an individual's stage of change [26].

Qualitative interviews identified a subgroup of participants who were satisfied with the intervention but did not report it being particularly impactful because of their already high levels of baseline health. This is in line with findings from our quantitative subgroup analysis from our recently completed randomized controlled trial, in that there was a significant effect of the intervention on physical activity levels of those with poor or fair baseline self-rated health. Measures of self-rated health have been found to be correlated with perceived physical fitness [27], physical mobility [28], and mortality [27,29]. Although baseline physical activity levels were controlled for in the analysis, those with the highest self-rated health at baseline may have been the most healthy and active; thus, change in physical activity throughout the intervention period may be limited by a ceiling effect, as suggested by several participants within the qualitative interviews. Interestingly, the percentage of participants who met physical activity guidelines was 27.4% in the intervention group and 29.4% in the control group at baseline. We hypothesize that individuals perhaps perceived they were sufficiently active and thus did not need to increase their physical activity levels, which contributed to the lack of change. Future interventions should consider including a feedback mechanism to bring awareness to participants' reported or measured physical activity levels prior to the start of the intervention.

A limitation to this process evaluation, and to our understanding of how engagement with the intervention materials correlated to behavior change, is our inability to access Google Analytics data on clicks per link from individual participants. We hypothesize that those who engaged most with the intervention materials may be most likely to have changed their behavior as a result of the intervention, however, we were not able to evaluate this, given the aggregate group-level data. Our sample

XSL•FO
RenderX

was primarily well-educated females; thus, our findings may be less applicable to online portals targeting other populations. In addition, all social media posts and hashtags were publicly available, so we are unable to attribute any tracked engagement exclusively to participants in our intervention group.

We have previously reported on the significant increase in physical activity levels observed in both groups and, in particular, the significant intervention effect in participants with low self-rated health at baseline. When combined with positive findings in this process evaluation on participant satisfaction and engagement with the intervention strategies and the Portal itself, we believe KT strategies such as those delivered through the Portal have the potential to be an effective, low-cost, and scalable intervention. Insights from this process evaluation suggest that targeting the appropriate population is an important consideration. Delivering the intervention to individuals with the greatest need (ie, those with low self-rated health) or to those with the greatest potential to show a change (ie, low baseline levels of physical activity) should be explored in future studies. Previous research has shown that individuals who use online health portals are typically more highly educated and have higher health literacy [30]. A challenge in conducting this work is to understand how to engage underserved groups in an online intervention such as ours. Further work is needed to understand which KT strategies may be most effective to increase knowledge, awareness, and engagement with a resource such as the Portal.

Acknowledgments

The authors would like to acknowledge Susannah Watson, Rawan Farran, and Rachel Warren for their assistance in completion of this study. The authors gratefully acknowledge the financial support from Suzanne Labarge via the Labarge Optimal Aging Initiative. This study was made possible by funding through the Labarge Opportunities Fund. SENS is supported by a postdoctoral fellowship from the Canadian Institutes of Health Research. The funding bodies had no role in the design of the study, data collection, analysis, or interpretation.

Conflicts of Interest

None declared.

References

1. Hirvensalo M, Rantanen T, Heikkinen E. Mobility difficulties and physical activity as predictors of mortality and loss of independence in the community-living older population. J Am Geriatr Soc 2000 May;48(5):493-498. [Medline: 10811541]
2. Bouaziz W, Vogel T, Schmitt E, Kaltenbach G, Geny B, Lang PO. Health benefits of aerobic training programs in adults aged 70 and over: A systematic review. Arch Gerontol Geriatr 2017;69:110-127. [doi: 10.1016/j.archger.2016.10.012] [Medline: 27912156]
3. Hardy SE, Perera S, Roumani YF, Chandler JM, Studenski SA. Improvement in usual gait speed predicts better survival in older adults. J Am Geriatr Soc 2007 Nov;55(11):1727-1734. [doi: 10.1111/j.1532-5415.2007.01413.x] [Medline: 17916121]
4. Hall KS, Cohen HJ, Pieper CF, Fillenbaum GG, Kraus WE, Huffman KM, et al. Physical performance across the adult life span: Correlates with age and physical activity. J Gerontol A Biol Sci Med Sci 2017 Apr 01;72(4):572-578 [FREE Full text] [doi: 10.1093/gerona/glw120] [Medline: 27356977]
5. Statistics Canada. 2013. Table 27-10-0016-01: Internet use by Internet activity, age group, sex, level of education and household income URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2710001601 [accessed 2019-06-11]
6. Silver MP. Patient perspectives on online health information and communication with doctors: A qualitative study of patients 50 years old and over. J Med Internet Res 2015 Jan 13;17(1):e19 [FREE Full text] [doi: 10.2196/jmir.3588] [Medline: 25586865]
7. Barbara AM, Dobbins M, Brian Haynes R, Iorio A, Lavis JN, Raina P, et al. McMaster Optimal Aging Portal: An evidence-based database for geriatrics-focused health professionals. BMC Res Notes 2017 Jul 11;10(1):271 [FREE Full text] [doi: 10.1186/s13104-017-2595-8] [Medline: 28693544]
8. McMaster Optimal Aging Portal. Hamilton, ON: McMaster University URL: https://www.mcmasteroptimalaging.org/ [accessed 2019-03-08] [WebCite Cache ID 76j9hfGCk]
9. Neil-Sztramko SE, Farran R, Watson S, Levinson AJ, Lavis JN, Iorio A, et al. If you build it, who will come? A description of user characteristics and experiences with the McMaster Optimal Aging Portal. Gerontol Geriatr Med 2017;3:1-9 [FREE Full text] [doi: 10.1177/2333721417737681] [Medline: 29152540]
10. Barbara AM, Dobbins M, Haynes RB, Iorio A, Lavis JN, Levinson AJ. User experiences of the McMaster Optimal Aging Portal's evidence summaries and blog posts: Usability study. JMIR Hum Factors 2016 Aug 19;3(2):e22 [FREE Full text] [doi: 10.2196/humanfactors.6208] [Medline: 27542995]
11. Barbara AM, Dobbins M, Haynes RB, Iorio A, Lavis JN, Raina P, et al. The McMaster Optimal Aging Portal: Usability evaluation of a unique evidence-based health information website. JMIR Hum Factors 2016 May 11;3(1):e14 [FREE Full text] [doi: 10.2196/humanfactors.4800] [Medline: 27170443]

XSL•FO
RenderX

12. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The impact of information technology on patient engagement and health behavior change: A systematic review of the literature. JMIR Med Inform 2016 Jan 21;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
13. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of eHealth interventions for the promotion of physical activity in older adults: A systematic review. Prev Med 2018 Dec;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]
14. Oakley A, Strange V, Bonell C, Allen E, Stephenson J, RIPPLE Study Team. Process evaluation in randomised controlled trials of complex interventions. BMJ 2006 Feb 18;332(7538):413-416 [FREE Full text] [doi: 10.1136/bmj.332.7538.413] [Medline: 16484270]
15. 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]
16. McMaster Optimal Aging Portal. 2016 Jan 28. Sitting too much? Take steps to support healthy aging URL: https://www.mcmasteroptimalaging.org/blog/detail/blog/2016/01/28/sitting-too-much-take-steps-to-support-healthy-aging [accessed 2019-06-04] [WebCite Cache ID 78spHU8YY]
17. Average industry rates for email as of March 2019. Waltham, MA: Constant Contact; 2019 May 07. URL: https://knowledgebase.constantcontact.com/articles/knowledgebase/5409-average-industry-rates?lang=en_US#mean [accessed 2019-06-05]
18. Song M, Ward J, Choi F, Nikoo M, Frank A, Shams F, et al. A process evaluation of a Web-based mental health portal (WalkAlong) using Google Analytics. JMIR Ment Health 2018 Aug 20;5(3):e50 [FREE Full text] [doi: 10.2196/mental.8594] [Medline: 30126832]
19. Boucher S, Edwards O, Gray A, Nada-Raja S, Lillis J, Tylka T, et al. Teaching intuitive eating and acceptance and commitment therapy skills via a Web-based intervention: A pilot single-arm intervention study. JMIR Res Protoc 2016 Oct 14;5(4):e180 [FREE Full text] [doi: 10.2196/resprot.5861] [Medline: 27742602]
20. Crutzen R, Roosjen JL, Poelman J. Using Google Analytics as a process evaluation method for Internet-delivered interventions: An example on sexual health. Health Promot Int 2013 Mar;28(1):36-42. [doi: 10.1093/heapro/das008] [Medline: 22377974]
21. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based intervention to promote physical activity by sedentary older adults: Randomized controlled trial. J Med Internet Res 2013 Feb 05;15(2):e19 [FREE Full text] [doi: 10.2196/jmir.2158] [Medline: 23470322]
22. Van Dyck D, Plaete J, Cardon G, Crombez G, De Bourdeaudhuij I. Effectiveness of the self-regulation eHealth intervention 'MyPlan1.0.' on physical activity levels of recently retired Belgian adults: A randomized controlled trial. Health Educ Res 2016 Dec;31(5):653-664. [doi: 10.1093/her/cyw036] [Medline: 27422898]
23. Wijsman CA, Westendorp RG, Verhagen EA, Catt M, Slagboom PE, de Craen AJ, et al. Effects of a Web-based intervention on physical activity and metabolism in older adults: Randomized controlled trial. J Med Internet Res 2013 Nov 06;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
24. King AC, Bickmore TW, Campero MI, Pruitt LA, Yin JL. Employing virtual advisors in preventive care for underserved communities: Results from the COMPASS study. J Health Commun 2013;18(12):1449-1464. [doi: 10.1080/10810730.2013.798374] [Medline: 23941610]
25. Crutzen R, de Nooijer J, Brouwer W, Oenema A, Brug J, de Vries NK. Strategies to facilitate exposure to Internet-delivered health behavior change interventions aimed at adolescents or young adults: A systematic review. Health Educ Behav 2011 Feb;38(1):49-62. [doi: 10.1177/1090198110372878] [Medline: 21189422]
26. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
27. Miilunpalo S, Vuori I, Oja P, Pasanen M, Urponen H. Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. J Clin Epidemiol 1997 May;50(5):517-528. [Medline: 9180644]
28. Singh-Manoux A, Martikainen P, Ferrie J, Zins M, Marmot M, Goldberg M. What does self-rated health measure? Results from the British Whitehall II and French Gazel cohort studies. J Epidemiol Community Health 2006 Apr;60(4):364-372 [FREE Full text] [doi: 10.1136/jech.2005.039883] [Medline: 16537356]
29. Idler EL, Benyamini Y. Self-rated health and mortality: A review of twenty-seven community studies. J Health Soc Behav 1997 Mar;38(1):21-37. [Medline: 9097506]
30. Coughlin SS, Prochaska JJ, Williams LB, Besenyi GM, Heboyan V, Goggans DS, et al. Patient Web portals, disease management, and primary prevention. Risk Manag Healthc Policy 2017;10:33-40 [FREE Full text] [doi: 10.2147/RMHP.S130431] [Medline: 28435342]

Abbreviations

KT: knowledge translation
Portal: McMaster Optimal Aging Portal

XSL•FO
RenderX

Edited by N Zary; submitted 08.03.19; peer-reviewed by M Silver, A Al-Zalabani; comments to author 12.04.19; revised version received 29.04.19; accepted 25.05.19; published 20.06.19.

Please cite as:
Neil-Sztramko SE, Smith-Turchyn J, Richardson J, Dobbins M
A Mobility-Focused Knowledge Translation Randomized Controlled Trial to Improve Physical Activity: Process Evaluation of the Move4Age Study
J Med Internet Res 2019;21(6):e13965
URL: http://www.jmir.org/2019/6/e13965/
doi:10.2196/13965
PMID:31223121

©Sarah E Neil-Sztramko, Jenna Smith-Turchyn, Julie Richardson, Maureen Dobbins. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Symptoms Targeted for Monitoring in a Web-Based Tracking Tool by Caregivers of People With Dementia and Agitation: Cross-Sectional Study

Kenneth Rockwood[1,2,3], MD; Myrlene Sanon Aigbogun[4], MPH; Justin Stanley[2], BScE; Helen Wong[2], MSc; Taylor Dunn[2], MSc; Chère A T Chapman[2], MBA; Susan E Howlett[2,3,5], PhD; Maia Miguelez[6], PhD; Lisa McGarrigle[1,2,3], PhD; Ross A Baker[4], PhD

[1]Geriatric Medicine Research Unit, Halifax, NS, Canada

[2]DGI Clinical Inc, Halifax, NS, Canada

[3]Division of Geriatric Medicine, Dalhousie University, Halifax, NS, Canada

[4]Otsuka Pharmaceutical Development & Commercialization Inc, Princeton, NJ, United States

[5]Department of Pharmacology, Dalhousie University, Halifax, NS, Canada

[6]Otsuka Canada Pharmaceutical Inc, Saint-Laurent, QC, Canada

Corresponding Author:
Kenneth Rockwood, MD
Geriatric Medicine Research Unit
Nova Scotia Health Authority
5955 Veterans' Memorial Lane
Halifax, NS, B3H 2E9
Canada
Phone: 1 902 473 8687
Fax: 1 902 473 1050
Email: kenneth.rockwood@dal.ca

Abstract

Background: In people with dementia, neuropsychiatric symptoms (NPSs), especially agitation, are associated with worse quality of life and caregiver burden. As NPSs may vary with illness severity, knowledge of how people with dementia and their caregivers describe and rate the importance of agitation symptoms can improve the understanding of the clinical meaningfulness of the manifestations of agitation. The internet provides new opportunities to better understand patient experiences, as patients and caregivers increasingly look to Web-based platforms as a means of managing symptoms.

Objective: The aim of this study was to examine Web-based reports from a dementia symptom website to better understand the symptoms of agitation and explore how they are being targeted for monitoring by caregivers of people with dementia.

Methods: The Dementia Guide website hosts a Web-based database used by caregivers (97%) and people with dementia (3%). From its 61 dementia symptoms, users can select relevant symptoms that they deem important to monitor or track the effects of treatment. We employed a staging algorithm to determine if individuals had mild cognitive impairment (MCI) or mild, moderate, or severe dementia. Agitation was defined using terms consistent with the International Psychogeriatrics Association's provisional consensus definition. We compared the proportion of people with NPSs and agitation across stages of dementia severity and studied how many agitation-defining descriptors were selected, and how often they occurred, by stage.

Results: As of March 2017, 4121 people had used the tracking tool, of whom 2577 provided sufficient data to allow disease severity staging. NPSs were tracked by 2127/2577 (82.54%) and agitation by 1898/2577 (73.65%). The proportion in whom agitation was tracked increased with increasing cognitive impairment: 68.5% (491/717) in people with MCI, and 72.50% (754/1040), 73.3% (378/516), and 90.5% (275/304) in mild, moderate, and severe dementia, respectively (χ^2_3=54.9; P<.001). The number of NPS and agitation descriptors selected also increased with severity (median number of NPSs=1, 2, 2, and 3 for MCI, mild, moderate, and severe dementia, respectively, Kruskal-Wallis H Test H_3=250.47; P<.001; median number of agitation descriptors=1, 2, 3, and 4, H_3=146.11; P<.001).

XSL•FO
RenderX

Conclusions: NPSs and agitation are common targets for tracking over the course of dementia and appear more frequently with increasing disease severity. These common and distressing symptoms represent clinically meaningful targets in treating people with dementia.

(J Med Internet Res 2019;21(6):e13360) doi:10.2196/13360

KEYWORDS

Alzheimer disease; dementia; agitation; neuropsychiatric symptoms; internet; caregiver

Introduction

Background

Cognitive decline is the hallmark of dementia, but more often the accompanying neuropsychiatric symptoms (NPSs) trouble people with dementia and their caregivers the most. NPSs appear across all types and stages of dementia [1-3] in as many as 80% to 90% of people with Alzheimer disease (AD) [4,5]. Apathy, irritability, agitation, depression, and anxiety typically are noted most often [6]. When specific NPSs arise, they vary with the type of dementia and its severity. In AD, apathy and depression are most often seen early, whereas delusions, hallucinations, and aggression become more prevalent with disease progression [2]. In its various guises, agitation is common across all stages of dementia [2]. This likely reflects agitation's multiple causes, including biological subsyndromes [7], reaction to declining cognitive function, and comorbid illnesses [8].

Defining Agitation

The understanding of agitation continues to evolve. Recent work to evaluate NPSs in dementia has focused on the symptom of agitation. Without a commonly accepted consensus description of agitation, it has been difficult to compare studies or even to know which behaviors are included in studies of agitation [9]. For these reasons, the International Psychogeriatric Association proposed a provisional consensus clinical and research definition, so as to help define populations for clinical care and research [9]. The proposed definition has resulted in a liberally construed notion of agitation, extending even beyond commonly used *broad-spectrum* NPS measures, such as the Neuropsychiatric Inventory (NPI) [10] and NPI-Clinician Rating Scale (NPI-C) [11]. In consequence of the newer approach, it is important to know how bothersome agitation-defining symptoms might be to people living with dementia and to their caregivers. As more bothersome symptoms vary with illness severity, knowledge of how caregivers describe and rate the importance of agitation symptoms can improve the understanding of the clinical meaningfulness of the myriad manifestations of agitation. One useful method for eliciting this information is by patient-centered, individualized measures, which allow patients (crucially in dementia) and their caregivers to specify which symptoms they find most troubling [12,13].

Web-Based Symptom Monitoring

The internet provides unprecedented access to the views of people living with chronic diseases. It is not yet clear as to how best their views can be elicited. Our group has pioneered one method of doing so. Since September 2006, the Dementia Guide website [14] offers a resource for persons with dementia (3% of users) and their caregivers (97% of users). Independently verified as credible using the DISCERN evaluation methodology [15], the site provides a mechanism for tracking symptoms important to caregivers and persons with dementia, thereby defining them as meaningful. This methodology is supported by previous work [16,17] and has allowed us to evaluate common but often understudied symptoms such as verbal repetition [18]. Although this approach cannot be used to estimate prevalence and incidence rates, which require representative population data, it can be used to compare frequencies in other nonrandom samples, and associations more generally, especially in relation to dementia severity.

Study Objectives

Here we used the dementia SymptomGuide to better understand the symptom of agitation and how it is being targeted for monitoring by caregivers. Specifically, our objectives were to (1) define agitation using the dementia SymptomGuide and evaluate how often it was being monitored across the stages of dementia severity, (2) estimate the degree of agitation experienced by people with dementia and how it varied with stage, and (3) compare the reported frequency (episodes per day) of agitation to other NPSs.

Methods

Ethics

The Research Ethics Committee at Nova Scotia Health Authority provided approval for this study. Approval was sought from Nova Scotia Health Authority as one of the authors on this study (KR) is affiliated with the institution. SymptomGuide users consented to terms of use, which included allowing their data to be aggregated and used for research purposes. Users were assured that research findings would be presented in a manner that would not disclose personal or individual identifying information.

Design, Participants, and Instrument

For this cross-sectional study, participants were recruited from the Dementia Guide website [14]. In addition to offering information, the website makes available the dementia SymptomGuide: a Web-based symptom tracker [19]. With the dementia SymptomGuide, individuals create a *Current Symptom Profile* to monitor the symptoms that are important to them. Users can choose from a standardized inventory of 61 symptoms and 609 symptom descriptions or enter symptoms and descriptions of their own. The predefined symptom descriptions overlap such that one symptom description can fall under more than one symptom. For example, the description *does not recognize current dwelling as home* from the *disorientation to place* symptom overlaps with *believes they live somewhere else*

XSL•FO
RenderX

from *delusions and paranoia*. This redundancy is a built-in feature to allow users multiple ways to address common phenomena. Users note the occurrence of each selected symptom and record its frequency at baseline and whenever they choose to follow up. SymptomGuide users are also asked to rank their chosen symptoms from most to least important (from most to least bothersome). Symptom rankings were normalized as the rank divided by total number of symptoms reported resulting in a 0 to 1 scale, such that a weighted rank of 1 is most important. At baseline, SymptomGuide users are also asked to provide information on demographics, dementia diagnosis, and medication use. For those who select 3 or more symptoms, the dementia stage can be determined using a staging algorithm [17]. In this way, users have an individualized selection of symptoms that they track with standardized descriptions to allow comparisons across users. In this study, we used data from the users' first visit (baseline) only.

We attempted to determine the level of cognitive impairment for all users with our staging algorithm [17]. Users who selected fewer than 3 symptoms were excluded from the analysis because the staging algorithm requires at least 3. Users for whom stage could be determined were categorized into 4 levels of cognitive impairment (mild cognitive impairment [MCI], mild dementia, moderate dementia, or severe dementia) as previously described by our group [17]. To control for outliers, if the number of symptoms selected by a user was greater than the 95th percentile (more than 23 symptoms), they were excluded from the analyses. Symptom frequencies were truncated at 10 episodes per hour (99th percentile) plus 1, such that symptom frequencies greater than 10 episodes per hour were imputed as 11 episodes per hour.

Mapping Agitation to Symptom Descriptions

Agitation itself does not appear as one of the 61 symptoms in the dementia SymptomGuide Library. Instead, descriptions of many symptoms refer to agitation and its key components. To operationalize agitation, we took terms consistent with the International Psychogeriatric Association and NPI-C agitation definition [9,11] and mapped each term to SymptomGuide symptom descriptions. For example, the term *screaming* from the International Psychogeriatric Association's consensus definition was mapped to the SymptomGuide descriptions: *yells, shouts, or screams*, *verbally attacks others (eg, shouts insults)*, *resists and refuses assistance with verbal outbursts*, and *yells at, or becomes easily angry with grandchildren*. In this way, 90 dementia SymptomGuide descriptions from 33 separate symptoms were combined to define agitation (Multimedia Appendix 1).

Statistical Analysis

Participant characteristics were reported as frequencies and percentages, means and SD, or medians and ranges as appropriate. Percentages were based on the number of individuals with available data for each specific characteristic.

Objectives were evaluated using Pearson chi-squared test to compare the reported proportion of agitation and NPSs between stages of dementia severity (objective 1), Kruskal-Wallis one-way analysis of variance by ranks to compare the number and proportion of NPSs and agitation descriptors across groups, and a post hoc analysis with the Mann-Whitney-Wilcoxon test to determine pairwise associations (objective 2), and anticipating a non-normal distribution, we analyzed differences in reported episodes per day between symptoms using the nonparametric Mann-Whitney-Wilcoxon test (objective 3). All analyses were performed using R statistical software (R Core Team, Austria). All post hoc multiple comparisons were performed using Bonferroni corrections, by multiplying *P* values by the number of comparisons. Differences were considered statistically significant for values less than .05.

Results

Design and Participants

As of March 2017, the dementia SymptomGuide database consisted of 4121 participants who created a symptom profile to monitor symptoms as targets for treatment. Of those, 2577 had 3 or more symptoms needed for the staging algorithm and so were included in the study (Figure 1). Most participants were elderly North American women (Table 1).

XSL•FO
RenderX

Figure 1. Participant inclusion chart. Participants were included in the analysis population if they met staging criteria excluding users who selected more than 23 symptoms (95th percentile). MCI: Mild Cognitive Impairment.

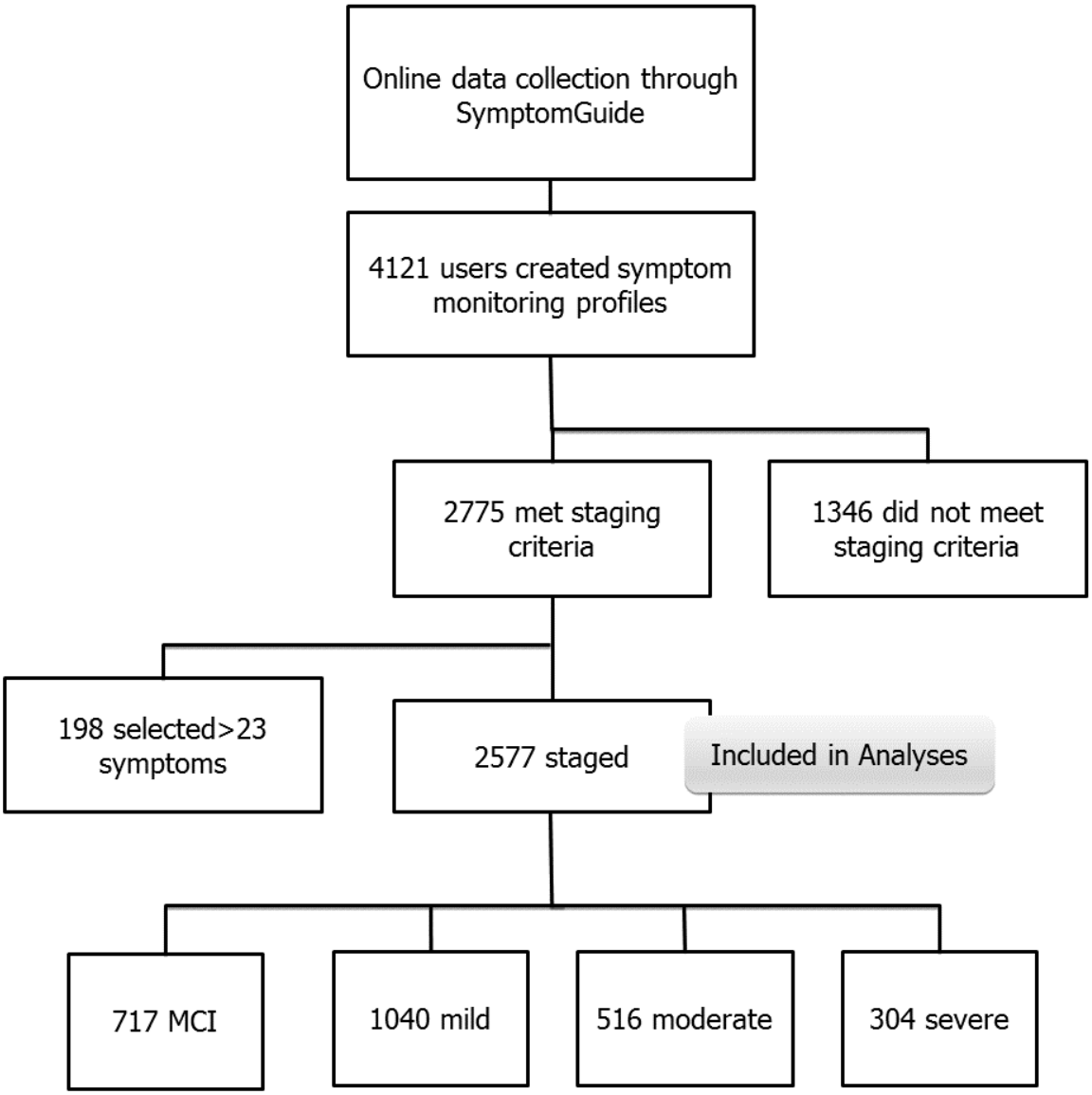


Table 1. Participant characteristics.

Characteristic	Total	Mild cognitive impairment	Mild	Moderate	Severe
Participants, n	2577	717	1040	516	304
Age (years), mean (SD[a])	76 (12)	74 (15)	75 (12)	77 (11)	77 (11)
Percentage[a] of women	63	63	65	61	62
Percentage[a] of participants with Alzheimer disease	61	100[b]	76	57	42
Percentage[a] living in care facility[c]	13	11	7	20	19
Percentage[a] with education ≥high school	77	82	78	78	71
Percentage[a] of North Americans	87	84	87	89	86

[a]Statistic/percentage of participants who reported information.

[b]Two participants reported disease type in the mild cognitive impairment group.

[c]Care facility included retirement homes and nursing homes.

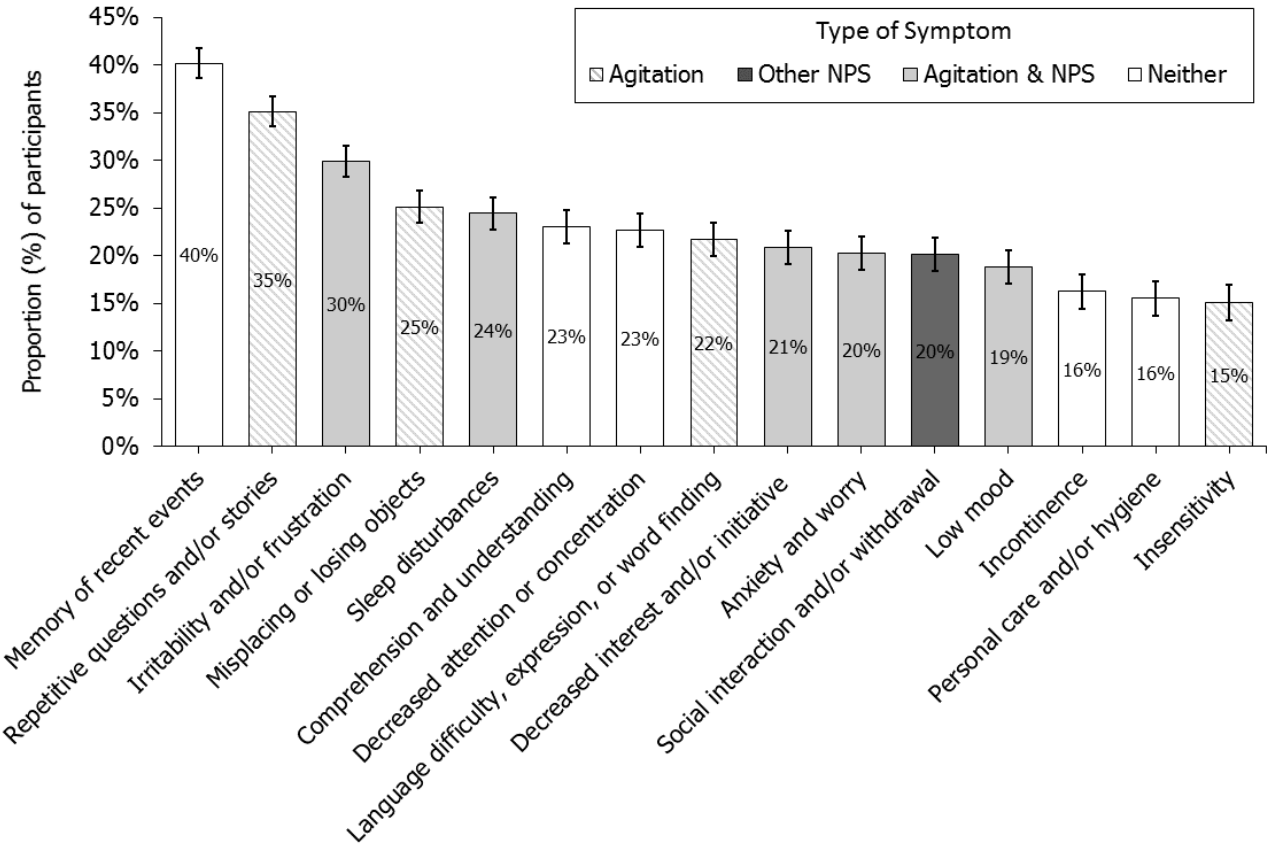
XSL•FO
RenderX

Reported Symptoms of Agitation

Of the 61 symptoms, 33 contained descriptions of agitation, of which 15 were NPSs (Multimedia Appendix 2). For example, one description of agitation from a non-NPS (*looking after grandchildren*) is *yells at or becomes easily angry with grandchildren*. Two-thirds of the most common symptoms (10/15) were an NPS, included a description of agitation, or were both (Figure 2). On average, 7.7 (SD 4.7) symptoms were targeted (median=4, range 3-23) per participant. In most participants (82.54%, 2127/2577), at least 1 symptom was an NPS. Agitation was selected across all stages with 73.65% (1898/2577) of participants reporting at least one description of the symptom. Agitation was reported least often in the MCI group (68.5%, 491/717) and most often in the severe dementia group (90.5%, 275/304).

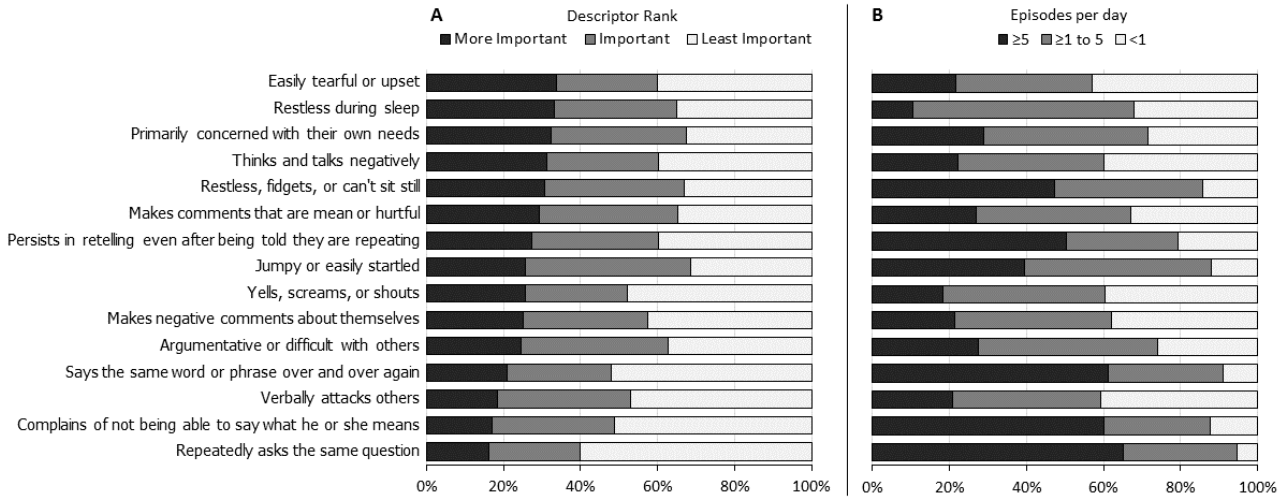
Agitation was most often described by verbalizations such as *Argumentative or difficult with others* (reported by 20.49%, 312/1523 of participants with ranked symptoms; Figure 3), *Makes comments that are mean or hurtful* (15.17%, 231/1523), *Yells, screams, or shouts* (8.93%, 136/1523), or *verbally attacks others* (8.54%, 130/1523). The 15 most commonly reported descriptions of agitation are shown in Figure 3. The frequency of the description did not correlate with the participants' ranking of importance of the symptom (r=.03; P=.1). For example, the symptom descriptions *verbally attacks others* and *says the same word or phrase over and over again* were ranked similarly (mean 0.42 for both) but had significantly different reported median episodes per day (1.6 versus 5.0; P=.01). The reported occurrence of any description of agitation was, on average, 7 times per day (median 2, interquartile range: 1.0-5.0) which was significantly more frequent than other NPSs at 4 times per day (median 1, interquartile range: 0.7-4.0; P<.001). The rankings of NPSs were rated more important than those for descriptions of agitation (median weighted rank 0.62 vs 0.51; P<.001).

Figure 2. Prevalence of symptoms most frequently being monitored on the dementia SymptomGuide by type of symptom. The proportion (% of participants) ±SEp monitoring the 15 most frequently monitored symptoms. "Agitation"=non-NPS including a description of agitation, "Other NPS"=NPS which does not include a description of agitation, "Agitation & NPS"=NPS that includes a description of agitation, "Neither"=symptom is not an NPS and does not include a description of agitation. NPS: Neuropsychiatric Symptoms, SEp: Standard Error of Proportion.



XSL•FO
RenderX

Figure 3. Ranking and daily frequency of the most common descriptions of agitation. Relationship between rank of importance (A) and number of episodes per day (B) of the 15 most commonly reported descriptions of agitation. "More Important" rank=1.00-0.67, "Important" rank=0.66-0.34, "Less Important" rank=0.33-0.00.



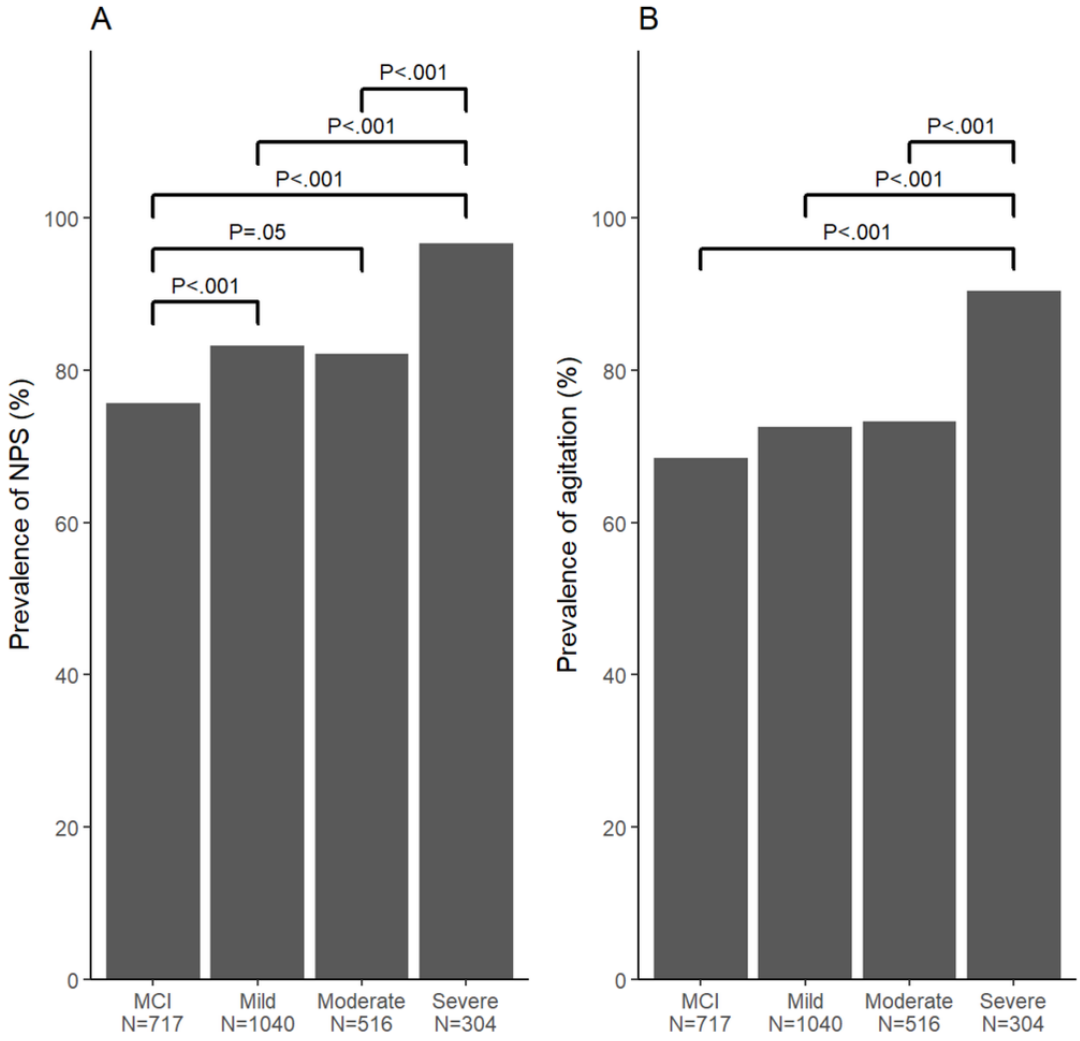
Variations by Severity

The proportion of participants with at least 1 NPS showed an increasing trend in relation to dementia severity—75.7% (543/717), 83.27% (866/1040), 82.2% (424/516), and 96.7% (294/304) for MCI, mild, moderate, and severe, respectively (χ^2_3=65.8; *P*<.001). Similarly, the proportion of participants with at least 1 agitation description increased with severity: 68.5% (491/717), 72.50% (754/1040), 73.3% (378/516), and 90.5% (275/304) for MCI, mild, moderate, and severe, respectively (χ^2_3=54.9; *P*<.001). A post hoc pairwise analysis showed that for people with severe dementia, there were high instances of both NPSs and agitation (Figure 4). We next compared the number of NPSs and agitation descriptions reported per participant by stage (Figure 5). There was a significant association between dementia severity and number of NPSs, with higher numbers of NPSs being reported in people with severe dementia (median 1, 2, 2, and 3 for MCI, mild, moderate, and severe, respectively; H_3=250.47; *P*<.001). Similarly, there was a significant association between severity and agitation, with higher numbers of agitation descriptors being reported in those with severe dementia (median 1, 2, 3, and 4; H_3=146.11; *P*<.001). For both measures, there were significant pairwise differences (*P*<.001) between every stage of dementia except mild and moderate. We also examined the prevalence of NPSs and agitation by stage and user-reported dementia diagnosis (Multimedia Appendix 3). The prevalence of both NPSs and agitation generally increased with dementia severity, in particular, in participants who reported a diagnosis of AD.

Finally, the proportion of total symptoms that were neuropsychiatric (number of NPSs divided by total number of symptoms) and the proportion of total descriptions related to agitation (number of agitation descriptions divided by total number of descriptions) were calculated for each participant (Figure 5). The proportion of NPSs varied significantly with stage, with the highest proportion occurring in those with severe dementia: median 27.3%, 25.0%, 25.0%, and 28.6% for MCI, mild, moderate, and severe, respectively (H_3=17.39; *P*<.001). Similarly, the proportion of agitation description was highest in those with severe dementia: median proportion 8.3%, 6.7%, 7.2%, and 9.2% (H_3=24.75; *P*<.001). Post hoc comparisons between stages revealed significant differences between the severe group and mild and moderate groups (*P*<.001 and *P*=.03, respectively) for NPSs and between the severe group and the MCI, mild, and moderate groups (*P*=.02, *P*<.001, and *P*=.002, respectively) for agitation. In participants who reported a specific dementia diagnosis, the number of NPSs and agitation descriptors selected increased with disease severity, significantly so in AD, vascular dementia, and frontotemporal dementia (Multimedia Appendix 4).

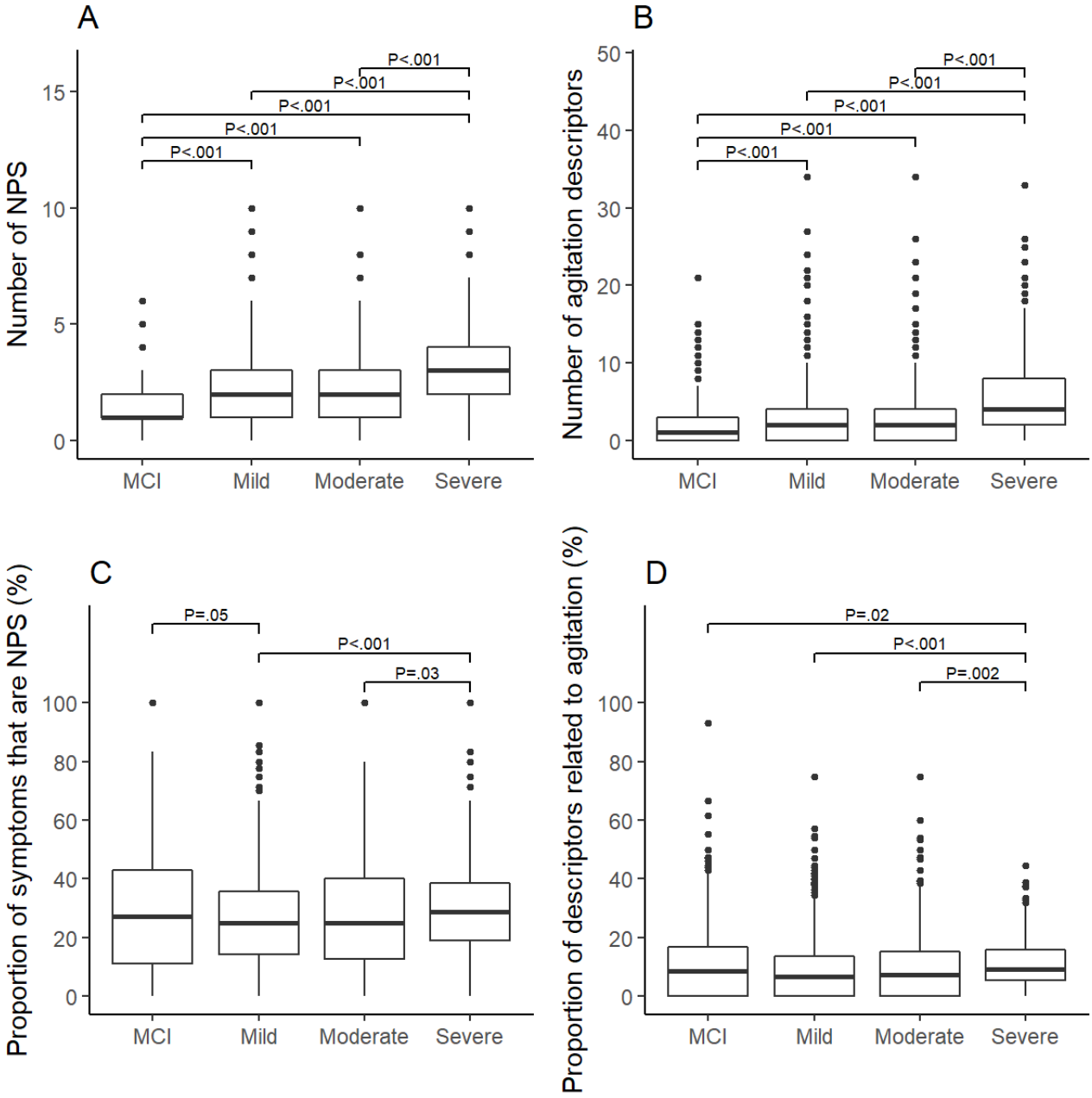
XSL•FO
RenderX

Figure 4. Proportion of participants who selected at least one NPS or description of agitation increased with stage. (A) % of participants who selected at least one NPS. (B) % of participants who selected at least one description of agitation. Bonferroni adjusted P values. MCI: Mild Cognitive Impairment, NPS: Neuropsychiatric Symptoms, p: P value.



XSL•FO
RenderX

Figure 5. Number and proportion of symptoms that were NPS or descriptions of agitation were higher in the severe stage. Number of NPS (A) and descriptors of agitation (B) reported per subject in each stage. Proportion of symptoms targeted that are NPS (C) and symptom descriptions targeted that are descriptions of agitation (D) by stage. Outliers are shown here as observations above Q3 + 1.5 IQR (the 75th percentile + 1.5 times the interquartile range). Bonferroni adjusted P values. IQR: Interquartile Range, MCI: Mild Cognitive Impairment, NPS: Neuropsychiatric Symptoms, p: P value.



Discussion

Principal Findings

Using the SymptomGuide Web-based tracking tool, we determined how dementia caregivers selected symptoms of agitation as targets for monitoring. Across all stages, symptoms that define agitation, including NPSs, were tracked by most users. The proportion in whom agitation was tracked increased with severity of cognitive impairment. Although most of the symptoms that were chosen for monitoring tended to occur frequently, a few were tracked even when they occurred uncommonly. This suggests that the impact of a given symptom to caregiver distress might reflect either its content or its frequency or both. The range of specific symptoms that were identified for tracking also suggests the merit of a broad-spectrum approach to agitation, and perhaps the need for individualization to make tracking feasible. As it is important to validate our methodological approach, we compared our estimates of frequency of individual NPSs and agitation symptoms (from Web-based patient/caregiver reports) to data collected more traditionally in memory clinics. A large, nonrandom study by Siafarikas et al [20] used data from the Norwegian national registry of memory clinics (n=4571) to examine the frequency of NPSs and NPS subgroups in MCI and the different stages of AD. The study thereby notably provides a recent and contemporary database of symptoms and associations, collected to the high standard inherent in a national registry. We were interested to know whether trends similar to those found there could be detected using our Web-based data

XSL•FO
RenderX

of predominantly North American caregiver reports. Recognizing that the differences in care and treatment practices between Norway and North America may be in play, nevertheless, our approach appears to be valid. For example, our estimate of the proportion of people with at least 1 NPS (82.54%, 2127/2577) corresponds closely to Norwegian estimates that NPSs were present in 87.2% of people with MCI or AD examined with validated neuropsychiatric cognitive assessments. Our data were consistent with that paper's report of at least 1 NPS in 79.5% of MCI subjects and in 91.2% of those with AD and showed a similar trend of increasing NPSs with increasing dementia severity (Figure 4) [20]. In further support of our findings, Peters et al [21] characterized NPSs using data from A Canadian Cohort Study of Cognitive Impairment and Related Dementias [22] and found similar trends, with 74% of those that were cognitively impaired-not demented and 89% of those with dementia reporting at least 1 NPS. In keeping with our findings, they also found that a greater proportion of subjects experienced agitation with increasing levels of cognitive impairment (23% in the CIND group vs 36% in the dementia group). Similar comparability was seen when we mapped the dementia SymptomGuide symptom descriptions to the Neuropsychiatric Inventory-Questionnaire (NPI-Q) agitation/aggression domain used in the Siafarikas et al [20] study. Note that in contrast to the 90 SymptomGuide descriptions that aligned with what was described in the International Psychogeriatric Association's provisional consensus report, only 50 such descriptions could be mapped to the NPI-Q agitation/aggression domain. However, the proportion of people with NPI-Q aggression/agitation increased with the severity of cognitive impairment—from 23% in people with MCI to 35% in people with moderate-severe dementia, close to our estimates using just those 50 descriptions: 19.0% (136/717) and 40.2% (330/820), respectively.

The similarity of our data with estimates from clinically adjudicated reports adds to our understanding of the validity and fidelity of Web-based data, at least where the Web-based reporting tool has been optimized for motivated users [19]. Resources such as the SymptomGuide increase the opportunity to capture the lived experience of people with dementia. In being able to demonstrate which of the symptoms with agitation are sufficiently salient to be monitored by caregivers, our data join those from the Norwegian registry report to offer empirical insights into the International Psychogeriatric Association's provisional consensus clinical and research definition.

We were also able to assess the degree to which individuals experienced NPSs and agitation. We found that the number of NPSs and agitation descriptions reported per participant increased with dementia severity and were significantly different between all stages except mild and moderate. This finding should be interpreted carefully, however, as the total number of reported symptoms (and by extension, the total number of descriptions) also generally increases with severity. More telling measures were the proportion of a participant's symptoms that were NPSs and the proportion of descriptions that related to agitation. We found similar estimates to those reported in other studies [20,21] showing that those with severe dementia experienced the highest degree of NPSs and agitation (median 28.6% of symptoms were NPSs; median 9.2% of descriptions were related to agitation). Of all the NPSs, agitation is often singled out as particularly troublesome and has become the focus of many investigations exploring the well-being of people with dementia and/or their caregivers [23,24]. To varying degrees of success, a range of pharmacological, psychological, and physical approaches are used to treat NPSs [1]. Despite their many advantages, instruments such as the NPI, NPI-Q, or NPI-C are highly structured and leave little room for customization or individualization by respondents. In other words, most self-report scales used for measuring NPSs contain predefined items or domains that have been deemed significant by the researcher, rather than the people with dementia and/or their caregivers. As a result, much less is known about the NPSs that are important to people with dementia or their caregivers, especially ones not included in NPS scales or measurements [9]. In this study, we did not have a comprehensive symptom inventory for each individual, however, we did have a list of symptoms selected for treatment, predominantly by caregivers. In this case, and as previously observed [25], agitation appears to be a potent symptom that is rarely ignored. The items selected by our caregiver/patient users were comparable in frequency and severity with those described by caregivers in the Norwegian national dementia registry, which also offers a comprehensive but nonrandom sample.

Limitations

Our data must be interpreted with caution. As the dementia SymptomGuide does not include a symptom for agitation, we used 90 of 609 specific symptom descriptions to define agitation. This wide range of symptoms is in keeping with what has been proposed in the provisional consensus definition of the International Psychogeriatric Association [9]. Arguably, focusing on descriptors of behavior versus interpretation of behaviors is a strength, particularly given that the participants are nonprofessionals. However, such a broad definition of agitation may over represent the number of cases in this sample. For example, in keeping with that proposal, we included verbal repetition in our definition of agitation. This also reflects observations laid out elsewhere that verbal repetition might reflect anxiety as much as impaired memory [18]. In this sense, verbal repetition could be considered a form of psychomotor agitation, a prospect that requires additional inquiry. Furthermore, we note that the data in this study consist in self-reported questionnaire responses, with nontechnical accounts completed by care providers and a few people with dementia. Even so, this approach is in line with the goal of developing clinically meaningful data [12]. In this study, only data from the first (baseline) Web visit were used. It would indeed be interesting to track the patterns of symptoms over time in future studies, as numbers accumulate to allow sufficient disaggregation of data. The introduction of mobile health apps for smartphones and tablets should accelerate data acquisition.

Conclusions

The importance of agitation symptoms as a target for treatment in our study supports the need for more well-controlled studies of both nonpharmacologic and pharmacologic treatments of agitation across all stages of AD, as well as the need for a

XSL•FO
RenderX

universally accepted operational definition of agitation in this context. In any era in which importance is being placed on the lived experience of people with dementia, how do we incorporate their point of view? How can we extricate their views from those of the people involved in their care? How can we balance the accessibility of Web-based resources with the interpretability of the data which might be generated on the Web? Can Web-based data be used to detect patterns in treatment response that might inform our understanding of mechanisms in dementia? These questions are motivating additional inquiries so that we can develop a more well-rounded understanding of how dementia affects the lives of the people to whom it is a challenge.

Acknowledgments

This study was conducted by DGI Clinical Inc, with funding provided by Otsuka Pharmaceuticals Development and Commercialization Inc.

Authors' Contributions

RAB, KR, CATC, SEH, MM, and MSA were involved in conceptualizing the study; RAB, KR, CATC, and MSA were responsible for the design of the study; KR was responsible for data acquisition; RAB, KR, TD, and JS participated in data analysis; RAB, KR, SEH, CATC, MM, MSA, TD, JS, HW, and LM were all involved in data interpretation, drafting, and critically reviewing the manuscript and also approved the manuscript for submission.

Conflicts of Interest

In addition to academic and hospital appointments, KR is President and Chief Science Officer of DGI Clinical, which in the last 5 years had contracts with pharma and device manufacturers (Baxter, Baxalta, Biogen, Shire, Hollister, Nutricia, Roche, and Otsuka) on individualized outcome measurement. In 2017, he attended an advisory board meeting with Lundbeck. Otherwise, any personal fees are for invited guest lectures, rounds and academic symposia, received directly from event organizers, for presentations on frailty. He is the Associate Director of the Canadian Consortium on Neurodegeneration in Aging, which is funded by the Canadian Institutes of Health Research (CNA-137794), with additional funding from the Alzheimer Society of Canada and several other charities, as well as from Pfizer Canada and Sanofi Canada (in phase 1, 2014-2019). He receives career support from the Dalhousie Medical Research Foundation as the Kathryn Allen Weldon Professor of Alzheimer Research and research support through grants from the Canadian Institutes of Health Research, the Canadian Frailty Network, the Nova Scotia Health Research Foundation, the Nova Scotia Health Authority Research Fund, and the Fountain Family Innovation Fund of the QEII Health Science Centre Foundation. RAB and MSA are employees of Otsuka Pharmaceutical Development and Commercialization Inc. MM is an employee of Otsuka Canada Pharmaceutical Inc. CATC is Chief Executive Officer of DGI Clinical. At the time of the study, HW, SEH, TD, and JS were employees of DGI Clinical Inc. LM received a postdoctoral fellowship from the Mitacs Elevate program, and her industrial cosponsor as part of this program is DGI Clinical Inc. Dementia Guide is a registered trademark owned by DementiaGuide Inc which operates as DGI Clinical.

Multimedia Appendix 1

Agitation descriptors.

[DOCX File, 20KB - jmir_v21i6e13360_app1.docx]

Multimedia Appendix 2

Dementia SymptomGuide symptom listing.

[DOCX File, 20KB - jmir_v21i6e13360_app2.docx]

Multimedia Appendix 3

Prevalence of neuropsychiatric symptom (NPS) and agitation increases with dementia stage, especially in Alzheimer disease. Proportion of participants who selected at least one NPS (A) or description of agitation (B), stratified by user-reported dementia type. Differences in proportions were tested in each dementia type by Pearson’s chi-squared test. Note that patients with MCI were not included in this analysis because they did not provide dementia diagnoses. Bonferroni adjusted P values.

[PNG File, 25KB - jmir_v21i6e13360_app3.png]

Multimedia Appendix 4

Number of neuropsychiatric symptom (NPS) and agitation descriptors increases with dementia stage in Alzheimer disease, vascular and frontotemporal dementia. The number of NPS (A) and descriptors of agitation (B) tracked per participant in each

XSL•FO
RenderX

stage, stratified by reported dementia type. Outliers were omitted for improved clarity. For each type of dementia, differences in number were tested by Kruskal-Wallis one-way ANOVA by ranks. Note that patients with mild cognitive impairment were not included in this analysis because they did not provide dementia diagnoses. Bonferroni adjusted P values.

[PNG File, 21KB - jmir_v21i6e13360_app4.png]

References

1. Livingston G, Sommerlad A, Orgeta V, Costafreda SG, Huntley J, Ames D, et al. Dementia prevention, intervention, and care. Lancet 2017 Dec 16;390(10113):2673-2674. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
2. Lyketsos CG, Carrillo MC, Ryan JM, Khachaturian AS, Trzepacz P, Amatniek J, et al. Neuropsychiatric symptoms in Alzheimer's disease. Alzheimers Dement 2011 Sep;7(5):532-539 [FREE Full text] [doi: 10.1016/j.jalz.2011.05.2410] [Medline: 21889116]
3. Steinberg M, Shao H, Zandi P, Lyketsos CG, Welsh-Bohmer KA, Norton MC, Cache County Investigators. Point and 5-year period prevalence of neuropsychiatric symptoms in dementia: the Cache County Study. Int J Geriatr Psychiatry 2008 Feb;23(2):170-177 [FREE Full text] [doi: 10.1002/gps.1858] [Medline: 17607801]
4. de Vugt ME, Riedijk SR, Aalten P, Tibben A, van Swieten JC, Verhey FRJ. Impact of behavioural problems on spousal caregivers: a comparison between Alzheimer's disease and frontotemporal dementia. Dement Geriatr Cogn Disord 2006;22(1):35-41. [doi: 10.1159/000093102] [Medline: 16679763]
5. Steinberg M, Tschanz JT, Corcoran C, Steffens DC, Norton MC, Lyketsos CG, et al. The persistence of neuropsychiatric symptoms in dementia: the Cache County Study. Int J Geriatr Psychiatry 2004 Jan;19(1):19-26. [doi: 10.1002/gps.1025] [Medline: 14716695]
6. Bergh S, Selbæk G. The prevalence and the course of neuropsychiatric symptoms in patients with dementia. Nor J Epidemiol 2012 Nov 29;22(2):-. [doi: 10.5324/nje.v22i2.1570]
7. Petrovic M, Hurt C, Collins D, Burns A, Camus V, Liperoti R, et al. Clustering of behavioural and psychological symptoms in dementia (BPSD): a European Alzheimer's disease consortium (EADC) study. Acta Clin Belg 2007;62(6):426-432. [doi: 10.1179/acb.2007.062] [Medline: 18351187]
8. Kales HC, Gitlin LN, Lyketsos CG, Detroit Expert Panel on Assessment Management of Neuropsychiatric Symptoms of Dementia. Management of neuropsychiatric symptoms of dementia in clinical settings: recommendations from a multidisciplinary expert panel. J Am Geriatr Soc 2014 Apr;62(4):762-769 [FREE Full text] [doi: 10.1111/jgs.12730] [Medline: 24635665]
9. Cummings J, Mintzer J, Brodaty H, Sano M, Banerjee S, Devanand DP, International Psychogeriatric Association. Agitation in cognitive disorders: International Psychogeriatric Association provisional consensus clinical and research definition. Int Psychogeriatr 2015 Jan;27(1):7-17 [FREE Full text] [doi: 10.1017/S1041610214001963] [Medline: 25311499]
10. Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology 1997 May;48(5 Suppl 6):S10-S16. [doi: 10.1212/WNL.48.5_Suppl_6.10S] [Medline: 9153155]
11. de Medeiros K, Robert P, Gauthier S, Stella F, Politis A, Leoutsakos J, et al. The Neuropsychiatric Inventory-Clinician rating scale (NPI-C): reliability and validity of a revised assessment of neuropsychiatric symptoms in dementia. Int Psychogeriatr 2010 Sep;22(6):984-994 [FREE Full text] [doi: 10.1017/S1041610210000876] [Medline: 20594384]
12. Reuben DB, Tinetti ME. Goal-oriented patient care--an alternative health outcomes paradigm. N Engl J Med 2012 Mar 1;366(9):777-779. [doi: 10.1056/NEJMp1113631] [Medline: 22375966]
13. Gordon MF, Lenderking WR, Duhig A, Chandler J, Lundy JJ, Miller DS, Patient-Reported Outcome Consortium's Cognition Working Group. Development of a patient-reported outcome instrument to assess complex activities of daily living and interpersonal functioning in persons with mild cognitive impairment: The qualitative research phase. Alzheimers Dement 2016 Jan;12(1):75-84. [doi: 10.1016/j.jalz.2015.04.008] [Medline: 26079412]
14. Dementia Guide. URL: https://www.dementiaguide.com [accessed 2019-05-13]
15. Dillon WA, Prorok JC, Seitz DP. Content and quality of information provided on canadian dementia websites. Can Geriatr J 2013;16(1):6-15 [FREE Full text] [doi: 10.5770/cgj.16.40] [Medline: 23440180]
16. Rockwood K, Zeng A, Leibman C, Mucha L, Mitnitski A. Validation of an informant-reported web-based data collection to assess dementia symptoms. J Med Internet Res 2012 Mar 12;14(2):e42 [FREE Full text] [doi: 10.2196/jmir.1941] [Medline: 22411293]
17. Rockwood K, Richard M, Leibman C, Mucha L, Mitnitski A. Staging dementia from symptom profiles on a care partner website. J Med Internet Res 2013 Aug 07;15(8):e145 [FREE Full text] [doi: 10.2196/jmir.2461] [Medline: 23924608]
18. Reeve E, Molin P, Hui A, Rockwood K. Exploration of verbal repetition in people with dementia using an online symptom-tracking tool. Int Psychogeriatr 2017 Dec;29(6):959-966 [FREE Full text] [doi: 10.1017/S1041610216002180] [Medline: 28274302]
19. Rockwood K. An individualized approach to tracking and treating Alzheimer's disease. Clin Pharmacol Ther 2010 Oct;88(4):446-449. [doi: 10.1038/clpt.2010.68] [Medline: 20856240]

XSL•FO
RenderX

20. Siafarikas N, Selbaek G, Fladby T, Šaltytė Benth J, Auning E, Aarsland D. Frequency and subgroups of neuropsychiatric symptoms in mild cognitive impairment and different stages of dementia in Alzheimer's disease. Int Psychogeriatr 2018 Dec;30(1):103-113. [doi: 10.1017/S1041610217001879] [Medline: 28927477]
21. Peters KR, Rockwood K, Black SE, Bouchard R, Gauthier S, Hogan D, et al. Characterizing neuropsychiatric symptoms in subjects referred to dementia clinics. Neurology 2006 Feb 28;66(4):523-528. [doi: 10.1212/01.wnl.0000198255.84842.06] [Medline: 16505306]
22. Feldman H, Levy AR, Hsiung G, Peters KR, Donald A, Black SE, ACCORD Study Group. A Canadian cohort study of cognitive impairment and related dementias (ACCORD): study methods and baseline results. Neuroepidemiology 2003;22(5):265-274. [doi: 10.1159/000071189] [Medline: 12902621]
23. Hongisto K, Hallikainen I, Selander T, Törmälehto S, Väätäinen S, Martikainen J, et al. Quality of Life in relation to neuropsychiatric symptoms in Alzheimer's disease: 5-year prospective ALSOVA cohort study. Int J Geriatr Psychiatry 2018 Jan;33(1):47-57. [doi: 10.1002/gps.4666] [Medline: 28067961]
24. Torrisi M, De Cola MC, Marra A, De Luca R, Bramanti P, Calabrò RS. Neuropsychiatric symptoms in dementia may predict caregiver burden: a Sicilian exploratory study. Psychogeriatrics 2017 Mar;17(2):103-107 [FREE Full text] [doi: 10.1111/psyg.12197] [Medline: 27411501]
25. Rockwood K, Fay S, Hamilton L, Ross E, Moorhouse P. Good days and bad days in dementia: a qualitative chart review of variable symptom expression. Int Psychogeriatr 2014 Aug;26(8):1239-1246. [doi: 10.1017/S1041610214000222] [Medline: 24565377]

Abbreviations

AD: Alzheimer disease
MCI: mild cognitive impairment
NPI: Neuropsychiatric Inventory
NPI-C: Neuropsychiatric Inventory-Clinician Rating Scale
NPI-Q: Neuropsychiatric Inventory-Questionnaire
NPS: neuropsychiatric symptom

Edited by G Eysenbach; submitted 15.01.19; peer-reviewed by Q Yuan, H McCarron, M O'Connell; comments to author 28.02.19; revised version received 09.04.19; accepted 26.04.19; published 28.06.19.

Please cite as:
Rockwood K, Sanon Aigbogun M, Stanley J, Wong H, Dunn T, Chapman CAT, Howlett SE, Miguelez M, McGarrigle L, Baker RA
The Symptoms Targeted for Monitoring in a Web-Based Tracking Tool by Caregivers of People With Dementia and Agitation: Cross-Sectional Study
J Med Internet Res 2019;21(6):e13360
URL: https://www.jmir.org/2019/6/e13360/
doi:10.2196/13360
PMID:31254339

©Kenneth Rockwood, Myrlene Sanon Aigbogun, Justin Stanley, Helen Wong, Taylor Dunn, Chère A T Chapman, Susan E Howlett, Maia Miguelez, Lisa McGarrigle, Ross A Baker. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Mobile Health Intervention for Prostate Biopsy Patients Reduces Appointment Cancellations: Cohort Study

Ashwin S Balakrishnan[1], MD; Hao G Nguyen[1,2], MD, PhD; Katsuto Shinohara[1,2], MD; Reuben Au Yeung[2], BSc; Peter R Carroll[1,2], MPH, MD; Anobel Y Odisho[1,2], MPH, MD

[1]Department of Urology, University of California San Francisco, San Francisco, CA, United States

[2]Helen Diller Family Comprehensive Cancer Center, San Francisco, CA, United States

Corresponding Author:
Anobel Y Odisho, MPH, MD
Department of Urology
University of California San Francisco
400 Parnassus Ave
San Francisco, CA,
United States
Phone: 1 415 476 1611
Email: anobel.odisho@ucsf.edu

Abstract

Background: Inadequate patient education and preparation for office-based procedures often leads to delayed care, poor patient satisfaction, and increased costs to the health care system. We developed and deployed a mobile health (mHealth) reminder and education program for patients scheduled for transrectal prostate biopsy.

Objective: We aimed to evaluate the impact of an mHealth reminder and education program on appointment cancellation rates, communication frequency, and patient satisfaction.

Methods: We developed a text message (SMS, short message service)–based program with seven reminders containing links to Web-based content and surveys sent over an 18-day period (14 days before through 3 days after prostate biopsy). Messages contained educational content, reminders, and readiness questionnaires. Demographic information, appointment cancellations or change data, and patient/provider communication events were collected for 6 months before and after launching the intervention. Patient satisfaction was evaluated in the postintervention cohort.

Results: The preintervention (n=473) and postintervention (n=359) cohorts were composed of men of similar median age and racial/ethnic distribution living a similar distance from clinic. The postintervention cohort had significantly fewer canceled or rescheduled appointments (33.8% vs 21.2%, *P*<.001) and fewer same-day cancellations (3.8% vs 0.5%, *P*<.001). There was a significant increase in preprocedural telephone calls (0.6 vs 0.8 calls per patient, *P*=.02) in the postintervention cohort, but not a detectable change in postprocedural calls. The mean satisfaction with the program was 4.5 out of 5 (SD 0.9).

Conclusions: An mHealth periprocedural outreach program significantly lowered appointment cancellation and rescheduling and was associated with high patient satisfaction scores with a slight increase in preprocedural telephone calls. This led to fewer underused procedure appointments and high patient satisfaction.

(J Med Internet Res 2019;21(6):e14094) doi:10.2196/14094

KEYWORDS

text messaging; appointments and schedules; mHealth; quality improvement; urology; prostate neoplasm

Introduction

Background

Surgical procedures are increasingly being performed in outpatient facilities such as physician offices or ambulatory care centers [1]. From 1981 to 2009, outpatient procedure volume grew from approximately 110,000 to 12 million procedures [2]. At least 60% of urologic procedures are performed in the outpatient setting, with prostate biopsy and cystoscopy being the most common outpatient urologic procedures [3]. Prostate biopsy remains essential for diagnosing and monitoring prostate cancer. However, safe in-office biopsy requires preprocedural preparation, such as prophylactic antibiotics, and postprocedure

XSL•FO
RenderX

symptom assessment. Inadequate education regarding the importance of prostate biopsy and preparation for the procedure can lead to delays in care, canceled appointments, decreased patient satisfaction, and costs to the health care system [4,5]. Prior reports have found that clinics can lose up to 16.4% of their daily potential revenue because of late cancellations or no-shows [6,7]. Moreover, patients often have anxiety about discomfort they may experience, procedure-related risks such as infection, and receiving concerning biopsy results, which likely contribute to missed appointments and nonadherence to patient instructions [8-10].

Health systems are expanding their use of mobile technologies to improve communication with patients. Several trials have shown that text message (short message service, SMS) reminders can improve medication adherence and clinical attendance for appointments [11,12]. Mobile health (mHealth) interventions have also led to improved rates of patient adherence in colon, breast, and cervical cancer surveillance cohorts [13-15]. In prostate cancer care, electronic health (eHealth) and mHealth interventions have helped patients understand their disease severity, weigh the risks and benefits of various treatment options, and track key information such as prostate-specific antigen (PSA) laboratory results [16]. However, there are no mHealth interventions in urology that address the actual receipt of care and the interaction between patients, providers, and the clinic [16]. Prostate biopsies are a relatively complex patient encounter that provides an opportunity to improve communication and efficient care delivery with an mHealth intervention. In order to undergo a safe prostate biopsy, patients may need to hold anticoagulation medications prior to the procedure, take prophylactic antibiotics at home, self-administer a rectal enema, or get preprocedural magnetic resonance imaging (MRI). After the procedure, patients are at risk for bleeding and infection.

Objectives

We developed and deployed an mHealth SMS-based reminder, education, and procedure preparedness assessment program for patients scheduled for transrectal prostate biopsy and evaluated the impact on patient appointment completion. We hypothesized that rates of canceled or rescheduled appointments would decrease following deployment of the program. We also evaluated communication frequency between patients and providers to assess for potential changes in provider workloads and patient satisfaction with the mHealth intervention following implementation of the program.

Methods

Development

We developed an SMS-based program for patients undergoing MRI–transrectal ultrasound (TRUS) fusion prostate biopsy at a busy academic urologic oncology practice at the University of California, San Francisco (UCSF). The program consisted of 8 text messages sent over an 18-day period, with the first message sent 14 days before prostate biopsy and the last message sent 3 days after the procedure (Table 1). The text messages contained short reminders or educational material with links to more detailed Web-based content and short questionnaires (2 to 4 questions) that could be viewed on a mobile phone. Educational content included step-by-step descriptions of the biopsy procedure, the importance of antibiotic and enema adherence, and an embedded animated video on the importance of getting a prostate biopsy. If patients had a mobile phone without internet capabilities, they could still view the text messages but did not have access to linked content. We developed software that integrated with the electronic health record (EHR; Epic Systems Corporation) to extract demographic data and contact information for a patient when they are scheduled for a prostate biopsy and automatically enroll them to receive text messages. We used services from a commercial provider (Medumo Inc) to send text messages and log results. Patients received a short message at the time of appointment scheduling which welcomed them to the program and allowed them to opt out of receiving messages.

We employed a development process detailed in an earlier report to prototype, refine, and evaluate the SMS-based intervention [17]. We first collated all materials given to patients in clinic or via the EHR patient portal and call scripts for periprocedural reminder phone calls. We engaged stakeholders (clinic managers, nurses, and urologists) to identify key prostate cancer concepts that patients expressed difficulty understanding and the nature and timing of preparatory steps that patients had difficulty following. With the guidance of clinical staff and providers, preparatory instructions and patient education concepts within all materials were identified, modified, and incorporated in the mHealth intervention. These included instructions and questionnaires for antibiotic adherence, anticoagulation management, enema use, and confirmation of completion of prostate MRI for patients undergoing MRI-fusion biopsy. Postprocedure symptoms were assessed via questionnaire. Messages and Web-based content were written to an 8th grade reading level using the Flesch-Kincaid Grade Level [18].

Any concerning preprocedure responses (such as failing to stop anticoagulation) triggered an email to the clinic manager, who then triaged follow-up to the appropriate nurse. Concerning postprocedure responses (fever, bleeding) prompted patients to contact the clinic (routed to the on-call physician after hours) in addition to triggering an email alert. Informed by the expertise of clinical providers and the American Urological Association guidelines for prostate biopsy, we identified the ideal temporality of content to guide when text messages should be sent to patients (Table 1) [19].

XSL•FO
RenderX

Table 1. Schedule of text messages sent to patients.

Day	Time	Content sent
At registration	Enrollment	Program welcome, patient homepage link
14 days before	10 am	Magnetic resonance imaging and medication survey
12 days before	10 am	Educational information and video on program
7 days before	9 am	Key items to obtain and fleet enema instructions
1 day before	9 am	Preprocedure readiness survey
Day of procedure	7 am	Antibiotic and Fleet enema reminder
Day of procedure	5 pm	Postprocedure precautions
2 days after	10 am	Follow-up symptom survey
4 days after	5 pm	Satisfaction survey

Study Design and Data Sources

The program was launched on May 1, 2018, as a practice-wide quality improvement initiative. The preintervention cohort was defined as patients undergoing prostate biopsy in the 6-month period prior to program launch (November 1, 2017, to April 30, 2018) and the postintervention cohort was defined as patients undergoing prostate biopsy during a 6-month period (June 1 to November 30, 2018); patients were required to have a phone number to a mobile phone with SMS capabilities stored in our medical record system. There was a 1-month washout period dividing the two cohorts to account for patients who had biopsies scheduled within 14 days and would therefore not receive the full sequence of text messages. We retrospectively obtained appointment and communication frequency statistics.

Demographic information (age, race/ethnicity, distance from home to our clinic [km], urban vs rural geography, Diez-Roux neighborhood score based on patient zip code as a proxy for socioeconomic status [20], and insurance type), data on the occurrence and timing of appointment cancellations and rescheduling, and patient-provider communications (patient-provider phone calls and Epic MyChart in-basket messages) were collected from the electronic medical record for patients in the preintervention and postintervention cohorts. Providers involved with patient communication included urologists, nurse practitioners, registered nurses, and nurse navigators.

Outcomes and Analyses

The primary outcome of this study was the percentage of prostate biopsy appointments that were canceled or rescheduled. We further categorized cancellations by their temporality (same-day or within 7 and 14 days of scheduled appointment). We defined cancellation lead time as the number of days before the scheduled appointment that the appointment was canceled or rescheduled. Secondary outcomes included nature and frequency of patient-provider communications and patient satisfaction. Communications were defined as being preprocedural (within 14 days before appointment), postprocedural (within 7 days after), or periprocedural (14 days before to 7 days after). In-basket messages were defined as all relevant patient-provider and provider-provider messages in the Epic MyChart portal. Patient messages to providers regarding medical advice, provider messages to patients, and provider-provider messages were also collected and categorized. Patient satisfaction, as gauged by three survey questions delivered via text message on the last day of the program, was evaluated in the postintervention cohort. The questions were scored on a 5-point Likert scale and included the following: (1) How highly would you recommend this digital instruction program to a family member or friend? (1 = would not recommend and 5 = would highly recommend), (2) Overall, how satisfied are you with the care you received? (1 = not satisfied at all and 5 = very satisfied), and (3) What did you think of the number of reminders? (1 = far too many messages, 3 = the right number of messages, and 5 = far too few messages).

Baseline differences in the preintervention and postintervention cohorts were compared using chi-square tests for categorical factors and sample *t* tests or Mann-Whitney tests for continuous factors. All analyses were performed using R 3.5.1 (The R Foundation). This study was approved by the UCSF institutional review board.

Results

Sample Characteristics

There were 473 patients in the preintervention cohort (November 1, 2017, to April 30, 2018) who did not receive the SMS program, and 359 patients in the postintervention cohort who were enrolled in the 18-day mHealth program (Table 1 and Figure 1). Four eligible patients (1.1%) in the postintervention cohort opted out of the mHealth program. The preintervention and postintervention cohorts were composed of patients of similar median age (67.0 vs 67.6 years, *P*=.55) and of comparable racial/ethnic demographics (75.3% vs 76.0% white, *P*=.44; Table 2). Patients in both cohorts lived a similar median distance from care (74 vs 73 km, *P*=.74) and primarily lived in urban or metropolitan areas (88.5% vs 87.1%, *P*=.68). There were no differences in socioeconomic status as measured by neighborhood score (*P*=.39).

XSL•FO
RenderX

Figure 1. Text message with link to Web-based questionnaire sent to patients the day before procedure.

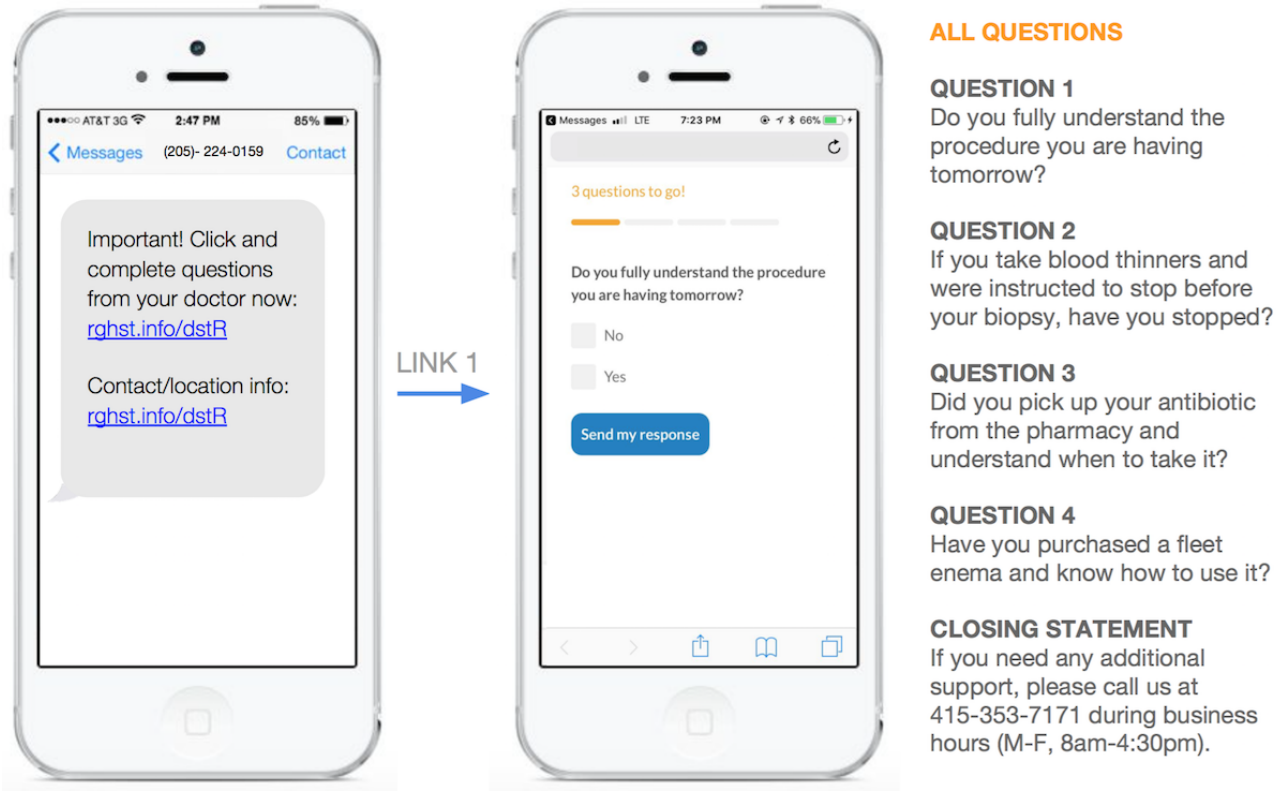


Table 2. Patient demographics in the preintervention and postintervention cohorts.

Characteristic	Preintervention (n=473)	Postintervention (n=359)	*P* value
Age in years, median (IQR[a])	67.5 (61.9-72.2)	67.0 (61.3-72.5)	.55
Ethnicity, n (%)			.44
Caucasian	356 (75.3)	273 (76.0)	
Black or African American	30 (6.4)	18 (5.0)	
Hispanic or Latino	14 (2.9)	17 (4.8)	
Asian	28 (5.9)	21 (6.0)	
Other/unknown	45 (9.6)	30 (8.3)	
Distance from clinic (km), median (IQR)	74 (21-202)	73 (23-192)	.74
Urban, n (%)	419 (88.5)	313 (87.1)	.68
Rural, n (%)	53 (11.2)	43 (12.1)	
Neighborhood score, median (IQR)	4.6 (1.7-7.5)	3.9 (1.3-7.5)	.39
Insurance, n (%)			.88
Commercial insurance	187 (39.6)	148 (41.2)	
Medicare	266 (56.3)	194 (54.0)	
Medi-Cal	18 (3.8)	15 (4.3)	
Self-pay	2 (0.3)	2 (0.5)	

[a]IQR: interquartile range.

Appointment Cancellation or Rescheduling

There were 37.3% fewer canceled or rescheduled appointments in the postintervention cohort compared with the preintervention cohort (33.8% vs 21.2%, *P*<.001; Table 3). Same-day cancellations were reduced by 86.8% with the intervention, with 3.8% of patients in preintervention cohort canceling on the day of their appointment compared with 0.5% in postintervention cohort (*P*<.001). Appointment cancellation or rescheduling within 7 days (13.1% vs 8.6%, *P*=.03) and within 14 days (19.0% vs 12.6%, *P*=.01) was also significantly lower in the postintervention cohort. There was not a detectable difference

XSL•FO
RenderX

in the median lead time to cancellation or rescheduling between the preintervention and postintervention cohorts (10.7 vs 10.8 days, *P*=.32). There were no detectable differences in patient age, race/ethnicity, geography, neighborhood socioeconomics, or insurance type between patients canceling or rescheduling appointments and those who did not.

Patient-Provider Communication

Compared with the preintervention period, in the postintervention period there was a significant increase in preprocedural in-basket messages (3.6 vs 4.1 messages per patient, *P*=.04) but not in postprocedural in-basket messages (1.3 vs 1.4 messages per patient, *P*=.56; Table 4). There were no detectable differences in periprocedural messages from patients to providers (1.6 vs 1.8, *P*=.34) or messages from providers to patients (1.5 vs 1.5, *P*=.88) when comparing the preintervention and postintervention cohorts. When looking at provider-provider in-basket message communication, there were significant increases in message volumes forwarding patient charts with comments (0.5 vs 0.9, *P*<.001) and clinic orders (0.6 vs 0.9, *P*<.006) in the postintervention period. After launching the intervention, there was an increase in preprocedural phone communication with patients (0.6 vs 0.8 telephone calls per patient, *P*=.02) but not in postprocedural phone communications.

Table 3. Appointment cancellation and rescheduling in the preintervention and postintervention cohorts.

Characteristic	Preintervention (n=627[a])	Postintervention (n=420[a])	*P* value
Total patients, n	473	359	
Appts[b], completed, n (%)	412 (65.7)	330 (78.6)	<.001
Appts, no-show, n (%)	3 (0.5)	1 (0.2)	.92
Appts, canceled or rescheduled, n (%)	212 (33.8)	89 (21.2)	<.001
Canceled or rescheduled, same-day	24 (3.8)	2 (0.5)	<.001
Canceled or rescheduled, within 7 days	85 (13.1)	38 (8.6)	.03
Canceled or rescheduled, within 14 days	119 (19.0)	54 (12.6)	.01
Cancel or reschedule lead days, median (IQR)[c]	10.67 (2.86-24.86)	10.79 (4.75-27.92)	.32

[a]All percentages are reported as a proportion of the total appointments scheduled in each column.

[b]Appts: appointments.

[c]IQR: interquartile range.

Table 4. Periprocedural communication volume per patient in preintervention and postintervention periods.

Characteristic	Preintervention, mean (SD)[a]	Postintervention, mean (SD)[a]	*P* value
All in-basket messages	4.9 (5.0)	5.6 (5.6)	.05
Preprocedural	3.6 (4.0)	4.1 (4.9)	.04
Postprocedural	1.3 (2.3)	1.4 (2.3)	.56
Patient message to provider	1.6 (2.6)	1.8 (3.0)	.34
Preprocedural	1.1 (2.0)	1.3 (2.4)	.11
Postprocedural	0.5 (1.3)	0.4 (1.2)	.46
Provider message to patient	1.5 (1.8)	1.5 (2.2)	.88
Preprocedural	1.2 (1.5)	1.2 (2.1)	.87
Postprocedural	0.3 (0.8)	0.3 (0.8)	.42
Provider-provider message	2.1 (2.4)	2.5 (2.2)	<.001
Preprocedural	1.5 (1.9)	1.7 (1.7)	.05
Postprocedural	0.6 (1.1)	0.8 (1.1)	.01
Patient-provider telephone call	0.8 (1.1)	1.0 (1.6)	<.001
Preprocedural	0.6 (1.0)	0.8 (1.5)	.02
Postprocedural	0.2 (0.5)	0.2 (0.5)	.24

[a]All values represent mean (SD) messages per patient, per appointment.

XSL•FO
RenderX

Patient Satisfaction

Mean patient satisfaction with the text message program was 4.5 out of 5 (SD 0.9), and the mean satisfaction with overall care was 4.8 out of 5 (SD 0.6, see Methods section for details on questions and scoring system). Patient opinion of text message quantity was 2.8 out 5, with a score of 3 corresponding to the right number of messages (SD 0.4).

Discussion

Principal Findings

An SMS-based mHealth periprocedural outreach program significantly lowered both last-minute and overall appointment cancellation and rescheduling and was associated with high patient satisfaction scores and a low opt-out rate. While the number of secure message and telephone interactions with patients slightly increased, this was associated with fewer underused procedure appointments and high patient satisfaction.

Preventing Appointment Cancellations

Appointment cancellations, particularly same-day cancellations and no-shows, can significantly burden health care systems. Additionally, inability to undergo the procedure after arrival due to inadequate preprocedure preparation is a significant time burden and inconvenience to patients, who often drive long distances from their home or take time off from work. For patients being evaluated for prostate cancer, procedure completion is essential for providing timely care. A meta-analysis of randomized controlled trials assessing the impact of SMS reminders in a wide range of practice settings found that reminders significantly increase attendance to health care appointments [12]. These results have been reinforced in the cancer screening literature. A systematic review of the impact of text message interventions on cancer screening rates found that absolute screening rates for patients receiving SMS reminders was 0.6% to 15% higher than for controls [14]. Cancer screening or surveillance patients sometimes share similar needs in terms of procedural preparation and patient education. For example, patients undergoing workup for colon cancer and prostate cancer both require adherence to several preparatory steps. In a randomized trial, Deng and colleagues found that an SMS reminder program significantly reduced cancellations from 8.0% to 4.8% in a clinic performing gastrointestinal endoscopy under sedation [13]. Similar to our program for patients undergoing prostate biopsy, their program contained reminders for key preparatory steps that often lead to cancellation, such as failure to discontinue an anticoagulant.

Selecting the Appropriate Platform

As more technologies for patient-provider communication become available, it is important to choose the appropriate method of communication to match the needs and technology literacies of specific patient populations. Web- and SMS-based technology has shown considerable promise in improving care for patients with prostate cancer. Kenfield and colleagues [21] found that patients with prostate cancer are amenable to using digital interventions (interactive website, text messaging, and a physical activity tracker) and that these interventions helped them adopt recommended lifestyle and dietary changes. A mobile phone app developed for detection and management of symptoms during prostate cancer treatment was found to reduce urinary-related symptoms and improve emotional functioning [22].

These earlier interventions for prostate cancer patients are designed for long-term care of patients while they undergo and recover from prostate cancer treatment. Patients are often highly invested in their preparation for surgery and therefore may be more willing to put in the effort to download a phone app or wear an activity tracking device. In comparison to surgery, prostate biopsy has a shorter periprocedural period and requires less physical and emotional preparation on the part of the patient. After considering several mobile app- or Web portal–based interventions, we decided to instead develop an SMS-based reminder program because this technology requires relatively less effort to engage with and has become widely adopted in our study population. Unlike mobile apps and Web portal–based interventions, SMS does not require the patient to download any software or create log-in credentials.

When deciding which patient communication platform to use, providers must assess the impact that any intervention will have on clinic workflow and capacity. We developed software in-house, which allowed us to automate enrollment in the text program. Practices that are not able to do this may need clinic staff to manually enroll patients in the program or rely on a commercial provider to help with this process. We also designed the program to trigger alerts to our clinic manager when a patient had a concerning response to a survey question. This also adds work for the clinical team. If a practice does not have capacity to manage these alerts, the survey questions and/or alerts can be removed. The program will still provide valuable reminders and educational information to patients without the triggered alerts.

Tailoring Interventions

Since reasons for cancellation often vary based on the type of appointment and patient population, customized interventions are needed to prevent cancellations and improve preparation [23]. Rather than simply reminding patients about their appointment, our intervention was customized to include educational information and survey questions specific to prostate biopsy patients. The study by O'Dwyer et al [23] of canceled elective urology appointments found that the majority of procedure-room cases (prostate biopsy, cystoscopy, and catheter changes) were canceled because patients were not adequately prepared for surgery. A study of procedural cancellations in a pediatric urology clinic found that many cancellations were due to preventable factors such as fasting violations [24]. These investigations reinforce the need for tailored interventions that address the specific needs of certain patient populations and complement the work of clinic staff rather than one-size-fits-all reminder systems.

Communication Volume

Patient communications with providers via telephone calls in the preprocedural period increased, while telephone calls in the postprocedural period did not significantly change. While the effect size of the preprocedural increase in call volume was

XSL•FO
RenderX

small, it is possible that it created meaningful changes in clinic workloads. Provider-provider communication also increased significantly, which may be due to increased communication about how to manage patient concerns reported over telephone or concerning responses to survey questions in the intervention. Increases in periprocedural communication may be necessary in order to adequately prepare patients for prostate biopsy, avoid scheduling inefficiencies, and prevent patient complaints that would lead to increased postprocedural communications.

Limitations

This study has limitations. Although we were able to track which patients clicked links in the text messages and opened Web-based content, for patients who did not click the links in the SMS, there is no other way with current SMS technology to assess if they received or read the message. This means that we are likely undercounting the degree of engagement, which biases our results toward the null. Patients were not directly engaged as stakeholders in the intervention development process. However, the program was based on materials given to patients in clinic that patients extensively helped to develop. We are currently gathering patient feedback on the program, which will directly inform future interventions. Although we aimed for an English 8th grade reading level, the intervention was not translated into other languages. Future interventions can assess patient language preferences and deliver language appropriate programs. The ethnic and racial composition of our cohort did not match the composition of the United States as a whole or the surrounding region, which may impact generalizability. Moreover, the majority of patients lived in urban areas, which likely impacts the generalizability of results to rural populations. Prior studies have found that SMS-based interventions have helped to improve clinical attendance and/or engage underserved populations in cancer screening and educational efforts [15,25,26]. Therefore, we aim to increase this proportion in future research by including clinical sites with greater proportions or racial/ethnic minorities and other vulnerable populations. However, there were no detectable demographic or socioeconomic differences between the preintervention and postintervention cohorts. Some patient-provider communications are not appropriately documented in the EHR. For example, communication with nonprovider clinic staff was not captured in our medical record and therefore not analyzed in this study. As an observational study, it is subject to selection bias. However, the study included six months of patient appointments in both the preintervention and postintervention cohorts, which may mitigate potential biases caused by short-term secular trends.

This study also has several strengths. First, we leveraged appointment scheduling logs in the EHR to obtain detailed information on the occurrence and timing of appointment cancellation or changes. Second, compared with other SMS-based interventions which require manual enrollment of patients, our intervention had an otherwise lower impact on clinic staff workloads as enrollment was automated and did not require any changes to workflow. Moreover, we were able to assess changes that the SMS program may have on clinic workloads by extracting and analyzing data on the type and frequency of patient-provider and provider-provider communications. Last, we considered the reading proficiency and technology literacy of our patient population in the design of the program.

Conclusions

An mHealth periprocedural outreach and reminder program designed specifically for prostate biopsy patients significantly lowered appointment cancellations and was associated with high patient satisfaction scores. The number of secure message and telephone calls per patient increased slightly; however, this increase in communication may be necessary in order to improve clinic efficiency and patient satisfaction. Future research on predictors of engagement with periprocedural SMS-based interventions and the impact of these programs with diverse study populations will help to understand the utility of this intervention among different patient groups.

Acknowledgments

The authors would like to sincerely thank Kayo Tsuruta for contributing her expertise in patient education and preparation for prostate biopsy to the design and development of this intervention.

Conflicts of Interest

None declared.

References

1. Urman RD, Punwani N, Shapiro FE. Office-based surgical and medical procedures: educational gaps. Ochsner J 2012;12(4):383-388 [FREE Full text] [Medline: 23267269]
2. Ohsfeldt RL, Li P, Schneider JE, Stojanovic I, Scheibling CM. Outcomes of surgeries performed in physician offices compared with ambulatory surgery centers and hospital outpatient departments in Florida. Health Serv Insights 2017;10:1178632917701025 [FREE Full text] [doi: 10.1177/1178632917701025] [Medline: 28469457]
3. Patel H, Matlaga B, Ziemba J. Trends in ambulatory urological surgery: an analysis of five states in the healthcare cost and utilization project. J Urology 2018 Apr;199(4S):e1022. [doi: 10.1016/j.juro.2018.02.2578]
4. Kheirkhah P, Feng Q, Travis LM, Tavakoli-Tabasi S, Sharafkhaneh A. Prevalence, predictors and economic consequences of no-shows. BMC Health Serv Res 2016 Jan 14;16(1):1-6 [FREE Full text] [doi: 10.1186/s12913-015-1243-z] [Medline: 26769153]

XSL•FO
RenderX

5. Dantas LF, Fleck JL, Cyrino Oliveira FL, Hamacher S. No-shows in appointment scheduling: a systematic literature review. Health Policy 2018 Dec;122(4):412-421. [doi: 10.1016/j.healthpol.2018.02.002] [Medline: 29482948]
6. Moore C, Wilson-Witherspoon P, Probst J. Time and money: effects of no-shows at a family practice residency clinic. Fam Med 2001;33(7):522-527. [Medline: 11456244]
7. Berg BP, Murr M, Chermak D, Woodall J, Pignone M, Sandler RS, et al. Estimating the cost of no-shows and evaluating the effects of mitigation strategies. Med Decis Making 2013 Nov;33(8):976-985 [FREE Full text] [doi: 10.1177/0272989X13478194] [Medline: 23515215]
8. Macefield RC, Metcalfe C, Lane JA, Donovan JL, Avery KNL, Blazeby JM, ProtecT Study Group. Impact of prostate cancer testing: an evaluation of the emotional consequences of a negative biopsy result. Br J Cancer 2010 Apr 27;102(9):1335-1340 [FREE Full text] [doi: 10.1038/sj.bjc.6605648] [Medline: 20372151]
9. Awsare NS, Green JSA, Aldwinckle B, Hanbury DC, Boustead GB, McNicholas TA. The measurement of psychological distress in men being investigated for the presence of prostate cancer. Prostate Cancer Prostatic Dis 2008;11(4):384-389. [doi: 10.1038/pcan.2008.21] [Medline: 18427569]
10. Chopra S, Rowe EWJ, Laniado M, Patel A. A prospective study analysing the effect of pain on probe insertion, and the biopsy strategy, on the patients' perception of pain during TRUS-guided biopsy of the prostate. N Z Med J 2008 Dec 12;121(1287):39-43. [Medline: 19098966]
11. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
12. Gurol-Urganci I, de Jongh T, Vodopivec-Jamsek V, Atun R, Car J. Mobile phone messaging reminders for attendance at healthcare appointments. Cochrane Database Syst Rev 2013;12:CD007458. [doi: 10.1002/14651858.CD007458.pub3] [Medline: 24310741]
13. Deng X, Wang Y, Zhu T, Zhang W, Yin Y, Ye L. Short message service (SMS) can enhance compliance and reduce cancellations in a sedation gastrointestinal endoscopy center: a prospective randomized controlled trial. J Med Syst 2015 Jan;39(1):169. [doi: 10.1007/s10916-014-0169-z] [Medline: 25476268]
14. Uy C, Lopez J, Trinh-Shevrin C, Kwon SC, Sherman SE, Liang PS. Text messaging interventions on cancer screening rates: a systematic review. J Med Internet Res 2017 Aug 24;19(8):e296 [FREE Full text] [doi: 10.2196/jmir.7893] [Medline: 28838885]
15. Muller CJ, Robinson RF, Smith JJ, Jernigan MA, Hiratsuka V, Dillard DA, et al. Text message reminders increased colorectal cancer screening in a randomized trial with Alaska Native and American Indian people. Cancer 2017 Apr 15;123(8):1382-1389. [doi: 10.1002/cncr.30499] [Medline: 28001304]
16. Pereira-Azevedo NM, Venderbos LDF. eHealth and mHealth in prostate cancer detection and active surveillance. Transl Androl Urol 2018 Feb;7(1):170-181 [FREE Full text] [doi: 10.21037/tau.2017.12.22] [Medline: 29594031]
17. Odisho AY, Gore JL. Patient-centered approaches to creating understandable health information. Urol Oncol 2017 Dec;35(9):559-563. [doi: 10.1016/j.urolonc.2017.06.055] [Medline: 28736251]
18. Ellimoottil C, Polcari A, Kadlec A, Gupta G. Readability of websites containing information about prostate cancer treatment options. J Urol 2012 Dec;188(6):2171-2175. [doi: 10.1016/j.juro.2012.07.105] [Medline: 23083852]
19. Liss MA, Ehdaie B, Loeb S, Meng MV, Raman JD, Spears V, et al. An update of the American Urological Association white paper on the prevention and treatment of the more common complications related to prostate biopsy. J Urol 2017 Dec;198(2):329-334. [doi: 10.1016/j.juro.2017.01.103] [Medline: 28363690]
20. Diez Roux AV, Merkin SS, Arnett D, Chambless L, Massing M, Nieto FJ, et al. Neighborhood of residence and incidence of coronary heart disease. N Engl J Med 2001 Jul 12;345(2):99-106. [doi: 10.1056/NEJM200107123450205] [Medline: 11450679]
21. Kenfield SA, Van Blarigan EL, Ameli N, Lavaki E, Cedars B, Paciorek AT, et al. Feasibility, acceptability, and behavioral outcomes from a technology-enhanced behavioral change intervention (Prostate 8): a pilot randomized controlled trial in men with prostate cancer. Eur Urol 2019 Jan 09:1-9. [doi: 10.1016/j.eururo.2018.12.040] [Medline: 30638635]
22. Langius-Eklöf A, Christiansen M, Lindström V, Blomberg K, Hälleberg Nyman M, Wengström Y, et al. Adherence to report and patient perception of an interactive app for managing symptoms during radiotherapy for prostate cancer: descriptive study of logged and interview data. JMIR Cancer 2017 Oct 31;3(2):e18. [doi: 10.2196/cancer.7599] [Medline: 29089290]
23. O'Dwyer A, Pridgeon S, Green J. Preventing 'same-day' cancellations in elective urological surgery: are different strategies needed for inpatient, day case and procedural admissions? J Clin Urol 2015 May 26;8(5):334-341. [doi: 10.1177/2051415815587213]
24. Pohlman GD, Staulcup SJ, Masterson RM, Vemulakonda VM. Contributing factors for cancellations of outpatient pediatric urology procedures: single center experience. J Urol 2012 Oct;188(4 Suppl):1634-1638. [doi: 10.1016/j.juro.2012.03.111] [Medline: 22910272]
25. Farmer T, Brook G, McSorley J, Murphy S, Mohamed A. Using short message service text reminders to reduce 'did not attend' rates in sexual health and HIV appointment clinics. Int J STD AIDS 2014 Mar;25(4):289-293. [doi: 10.1177/0956462413502325] [Medline: 23999939]

XSL•FO
RenderX

26. Le D, Holt C, Saunders D, Wang M, Coriolan A, Savoy A, et al. Feasibility and acceptability of SMS text messaging in a prostate cancer educational intervention for African American men. Health Informatics J 2016 Dec;22(4):932-947. [doi: 10.1177/1460458215598636] [Medline: 26324051]

Abbreviations

EHR: electronic health record
IQR: interquartile range
MRI: magnetic resonance imaging
PSA: prostate-specific antigen
SMS: short message service
TRUS: transrectal ultrasound
UCSF: University of California, San Francisco

Edited by G Eysenbach; submitted 21.03.19; peer-reviewed by C Ellimoottil, K Schaefer; comments to author 17.04.19; revised version received 03.05.19; accepted 05.05.19; published 02.06.19.

Please cite as:
Balakrishnan AS, Nguyen HG, Shinohara K, Au Yeung R, Carroll PR, Odisho AY
A Mobile Health Intervention for Prostate Biopsy Patients Reduces Appointment Cancellations: Cohort Study
J Med Internet Res 2019;21(6):e14094
URL: https://www.jmir.org/2019/6/e14094/
doi:10.2196/14094
PMID:31199294

©Ashwin S Balakrishnan, Hao G Nguyen, Katsuto Shinohara, Reuben Au Yeung, Peter R Carroll, Anobel Y Odisho. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 02.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Mobile Apps for Increasing Treatment Adherence: Systematic Review

Virtudes Pérez-Jover[1*], PhD; Marina Sala-González[1*], BSc; Mercedes Guilabert[1*], PhD; José Joaquín Mira[1*], PhD

Departamento Psicología de la Salud, Universidad Miguel Hernández, Elche, Spain

[*]all authors contributed equally

Corresponding Author:
Mercedes Guilabert, PhD
Departamento Psicología de la Salud
Universidad Miguel Hernández
Altamira Building, Avda de la Universidad s/n
Elche, 03202
Spain
Phone: 34 966658600 ext 8984
Fax: 34 966658984
Email: mguilabert@umh.es

Abstract

Background: It is estimated that 20% to 50% of patients do not take their medication correctly, and this leads to increased morbidity and inefficacy of therapeutic approaches. Fostering treatment adherence is a priority objective for all health systems. The growth of mobile apps to facilitate therapeutic adherence has significantly increased in recent years. However, the effectiveness of the apps for this purpose has not been evaluated.

Objective: This study aimed to analyze whether mobile apps are perceived as useful for managing medication at home and if they actually contribute to increasing treatment adherence in patients.

Methods: We carried out a systematic review of research published using Scopus, Cochrane Library, ProQuest, and MEDLINE databases and analyzed the information about their contribution to increasing therapeutic adherence and the perceived usefulness of mobile apps. This review examined studies published between 2000 and 2017.

Results: Overall, 11 studies fulfilled the inclusion criteria. The sample sizes of these studies varied between 16 and 99 participants. In addition, 7 studies confirmed that the mobile app increased treatment adherence. In 5 of them, the before and after adherence measures suggested significant statistical improvements, when comparing self-reported adherence and missed dose with a percentage increase ranging between 7% and 40%. The users found mobile apps easy to use and useful for managing their medication. The patients were mostly satisfied with their use, with an average score of 8.1 out of 10.

Conclusions: The use of mobile apps helps increase treatment adherence, and they are an appropriate method for managing medication at home.

(J Med Internet Res 2019;21(6):e12505) doi:10.2196/12505

KEYWORDS

mobile health; medication alert systems; medication adherence

Introduction

Background

The World Health Organization has classified the lack of treatment adherence as a major global problem [1]. This is partly because of therapeutic nonadherence being associated with high health costs because of rehospitalizations as a consequence of the lack of therapeutic response, changes in prescriptions for other more potent and toxic medications that increase the risk of producing side effects or long-term medication dependence, and, above all, the decreased efficacy from medication that patients either do not take or take inappropriately [2,3]. These consequences lead to increased morbidity and mortality in nonadherent patients [2,4].

Figures for Therapeutic Nonadherence

It is estimated that 20% to 50% of patients do not take their medication correctly [5-7]. The reasons for this lack of adherence to treatments are varied. On the one hand, patients

XSL•FO
RenderX

may voluntarily stop taking their medication because of, for example, a perception of the lack of improvements, beliefs that they have not been diagnosed correctly, or the adverse effects of the drug. However, the most frequent reasons for therapeutic nonadherence are involuntary causes, such as confusion or simple forgetfulness [2,5,8,9]. Medication errors at home are more usual than expected by health care professionals [8,10,11]. This would seem to indicate that designing and applying methods that foster treatment adherence and contribute to reduce medication errors at home are necessary.

Pillbox

The most commonly used device to promote medication adherence is the pillbox. People can independently manage their medications, check whether they have taken them or not, avoid the risk of taking them twice or not taking them at all, and reduce the rate of medication errors. Previous studies found that people who used a pillbox had better treatment adherence [12-14]. There are Medication Event Monitoring Systems (MEMSs), whose popular name is electronic pillbox. They have the additional feature of reminding patients to take the medication with alarms, and they are considered as the gold standard for measuring adherence [15]. However, unfortunately, none of these pillboxes are exempt from problems. They are too big to get them out of the house. In addition, patients have to understand the prescribed therapeutic regimen to organize the medication in the compartments and know how to manage these pillboxes [12-17].

Smartphones and Health Apps

All the research indicates that new information technologies have been rapidly accepted by the entire population [18]. In the case of Spain, 94.6% of its population currently uses a mobile phone [19].

This boom in mobile phones has resulted in these devices being used to devise new procedures to promote therapeutic adherence. At first, short message service (SMS) text messages were sent and telephone calls were made to remind users of the need to take medication. These kinds of reminders have been very effective methods and are well accepted by patients [20-22].

Then, with the advent of smartphones came mobile apps that have also afforded new opportunities for carrying out actions that simplify daily tasks, among them caring for health [6,8,18,23-25]. Currently, there are more than 165,000 apps designed for these devices that are related to health, and one in 5 people have downloaded a mobile health (mHealth) app [18]. Among these apps are a growing number intended to help patients in the management of their disease and their medication, remind users to take their drugs, and provide them with information about how they should do it to promote treatment adherence. These mobile apps are not only intended to help people remember to take the medication, such as the electronic pillbox; they have additional useful features that not only promote medication adherence but also increase treatment adherence.

However, very little research has been undertaken to evaluate the effectiveness of these apps for the purposes for which they were intended or the level of acceptance among users [6,18,26]. There are also no studies about their contribution to safe medication use.

Objective of This Study

This study aimed to analyze whether mobile apps that help people manage their medication in the home contribute to increasing patient adherence and are considered useful by the users.

Methods

A systematic review study that applied the recommendations in the Preferred Reporting Items for Systematic reviews and Meta-Analysis declaration for these types of studies was carried out [27].

Concepts to Be Taken Into Account in This Review

Pillbox

A small container that pills are carried in. A pillbox can make the medication task easier because it helps people to manage their daily medication. This device is associated with improvements in medication adherence and, subsequently, with better health [12].

Electronic Pillbox

The MEMS is a pill organizer that has the additional feature of reminding you to take your drugs with visual and audio alerts. This MEMS provides information about treatment adherence. Therefore, it is the gold standard for this purpose [17].

Mobile Apps

Mobile apps are computer programs or software installed on mobile electronic devices that supports a wide range of functions and uses, including television, telephone, video, music, word processing, and internet services [16].

Mobile Apps to Improve Medication Adherence

In this study, we considered the kind of mobile apps that help people to manage their medication. These mobile apps, compared with pillboxes or electronic pillboxes, have the main advantage of being a system that is incorporated into our smartphones [16].

Selection of Studies

The inclusion criteria for this review included research published in either English or Spanish that provided results about the effectiveness or treatment adherence in using mobile apps in the management of medication in the home, with any age group as the study population and regardless of the pathology and prescribed medication. Both quantitative and qualitative research were included, as well as research with descriptive and experimental approaches. The studies included in this review included presentations of results about the effectiveness in fostering adherence to treatment, safe medication use, viability, acceptance, satisfaction, and usefulness of these mobile apps. We excluded studies that were merely descriptive about the design of the mobile apps without presenting the results of use experience. We also excluded studies in which the interventions to remind patients to take their medication were delivered via SMS text messaging, phone calls, or electronic pillboxes.

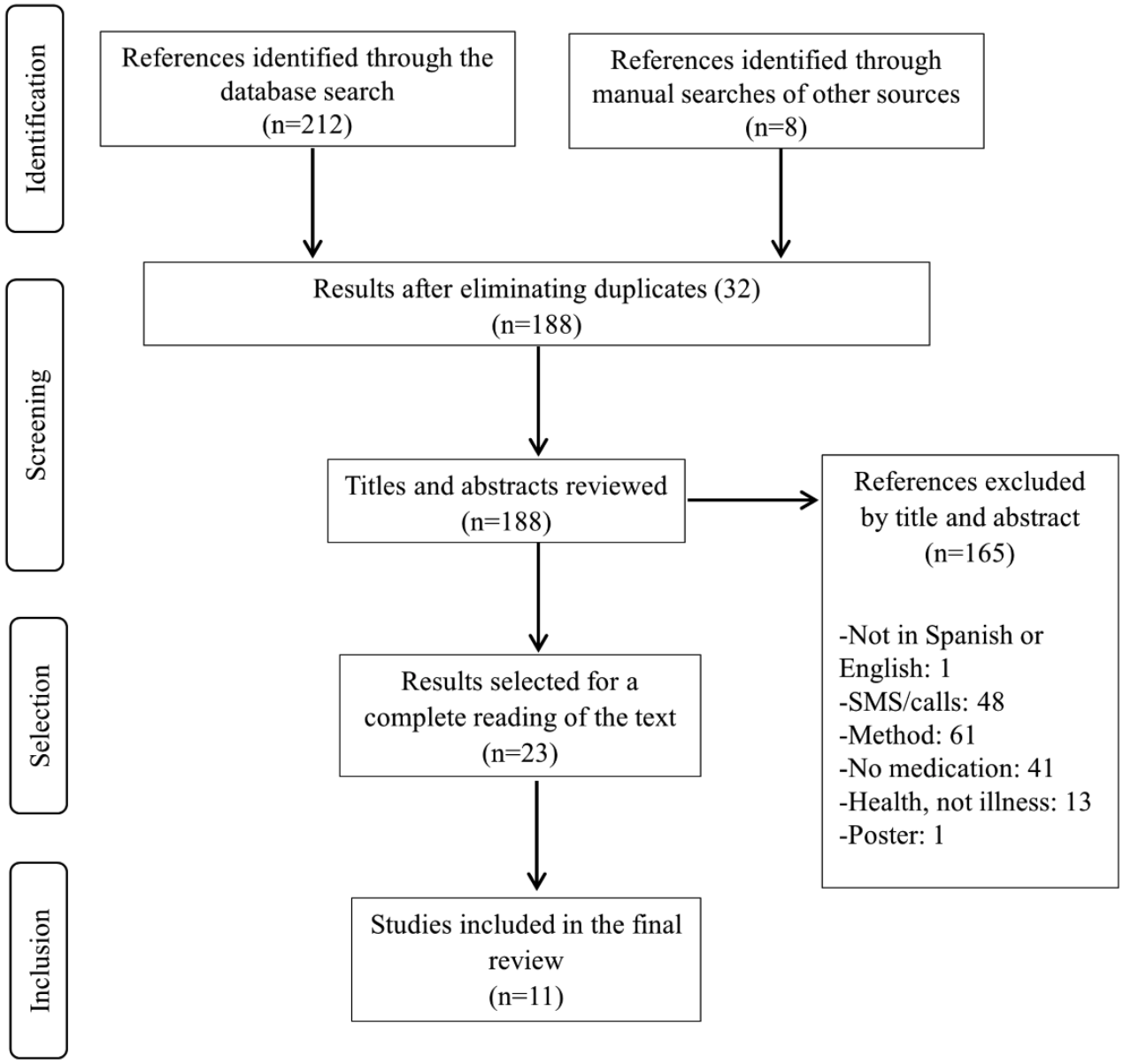
XSL•FO
RenderX

Search Strategy

We carried out a search for scientific documentation in the Scopus, Cochrane Library, ProQuest, and MEDLINE databases using keywords associated with pillbox and mobile apps and using the Boolean indicators OR and AND (*pillbox* OR *pill reminder* OR *pill organizer* OR *pill dispenser* OR *medication organizer* OR *medication reminder* OR *medication systems* OR *medicine reminder* OR *reminder system* AND *mhealth* OR *mobile app* OR *mobile application*). The search for documents was limited to publications that appeared in scientific journals from January 2000 through January 2017. The same descriptors were used to search the internet for relevant gray literature using the Google search engine. We similarly undertook a manual search using the bibliographic references of the selected publications.

The initial search identified 212 papers, of which 32 were eliminated because of being duplicates. Similarly, we found an additional 8 studies within either the bibliographies of the articles selected or through a Google search. We analyzed the titles and abstracts and eliminated 188 papers because they did not fulfill the inclusion criteria. We then fully read the 23 remaining papers and discarded 10 of them because they did not evaluate the effectiveness of the mobile app. Ultimately, 11 papers fulfilled the inclusion criteria (Figure 1).

Figure 1. Flow diagram of the study inclusion and exclusion process.



Data Extraction

The data extracted for each study included its country, objective, participants, chronic condition, design, and duration. Moreover, we recorded the functions of each mobile app, names of its designer(s), measures for evaluating adherence, measures for evaluating the mobile app, and outcomes of its evaluation.

In addition, to evaluate the quality of the reviewed publications, we first analyzed the level and degree of evidence following the classification proposed by the Scottish Intercollegiate

XSL•FO
RenderX

Guidelines Network [28]. Then, we assessed the following criteria (with a dichotomous yes/no scale): if it was reflected in the study that patients had participated to some degree in the app design, if the sample error had been controlled by adjusting the size of the sample under study, if there had been randomization with the samples who participated in the study to determine the app effectiveness, if validated measuring scales had been used, and whether the app had been used under natural conditions for periods of time exceeding 3 months.

We classified the levels of evidence as follows: 1++ (meta-analyses, systematic reviews of clinical trials, or high-quality clinical trials with very little risk of bias), 1+ (meta-analyses, systematic reviews of clinical trials, or well-conducted clinical trials with little risk of bias), 1− (meta-analyses, systematic reviews of clinical trials, or clinical trials with high risk of bias), 2++ (systematic reviews of cohort or case-control studies or studies of high-quality diagnostic tests and cohort or case-control studies of high-quality diagnostic tests with very little risk of bias and a high probability of establishing a causal relationship), 2+ (cohort or case-control studies or studies of well-conducted diagnostic tests with a low risk of bias and a moderate probability of establishing a causal relationship), 2− (cohort or case-control studies with a high risk of bias), 3 (nonanalytical studies, such as case reports and case series), and 4 (expert opinions) [28].

We classified the strengths of the recommendations as (A) at least 1 meta-analysis or systematic review of a controlled and randomized trial (CRT) or a level 1++ CRT, directly applicable to the target population or sufficient evidence extrapolated from 1+ level studies, directly applicable to the target population and whose results demonstrate overall consistency; (B) sufficient evidence deriving from level 2++ studies, directly applicable to the target population and whose results demonstrate overall consistency, with evidence extrapolated from either 1++ or 1+ level studies; (C) sufficient evidence deriving from level 2+ studies, directly applicable to the target population and whose results demonstrate overall consistency, with evidence extrapolated from level 2++ studies; and (D) evidence from either level 3 or 4 studies, with evidence extrapolated from level 2+ studies [28].

The evaluation and classification of the studies found during the search strategy were completed independently by 2 investigators (MS and VPJ). Discordant elements were discussed by both investigators until an agreement was reached.

Results

The initial search identified 212 papers. Ultimately, 11 papers fulfilled the inclusion criteria (Figure 1). Multimedia Appendix 1 shows the level of evidence and the degrees of recommendation of each of the 11 selected studies. Multimedia Appendix 2 shows the assessment of the internal quality of the design of each study.

Study Objectives

In 7 of the 11 studies [9,18,24,29-32] analyzed, a mobile app was designed and evaluated, whereas 4 studies [23,33-35] evaluated a previously designed app. Furthermore, 7 studies [9,23,29-31,33,35] evaluated both the perceived usefulness and treatment adherence when using a mobile app. In addition, 4 studies assessed perceived usefulness but did not evaluate adherence [18,24,32,35].

Moreover, 1 study [34], in addition to evaluating the mobile app's viability and acceptance, compared the intervention of 4 groups: mobile app with a reminder, mobile app without a reminder, electronic pillbox with a reminder, and electronic pillbox without a reminder. The objective of another study [18] was to evaluate the ad hoc–designed mobile app and compare the responses between people both older and younger than 55 years. Yet another study [35] compared the ease of use and usefulness of various mobile apps found for managing medication.

Description of the Population in the Reviewed Studies Using These Mobile Apps

The sample sizes varied between 16 and 99 participants [9,24]. The participants' ages varied depending upon the study. Of the studies, 2 focused on adolescents [24,29], 1 was directed at persons over the age of 65 years [9], another 1 was for persons over the age of 60 years [31], and another 1 for persons over the age of 50 years [35]. In addition, one study included adults with a wide age range (from 45-90 years) [34] and another compared the responses between people older and younger than 50 years [18].

The investigators recruited the samples at hospitals [23,24,29,30,33], health centers [9,32], patient associations [23,32,34], and local cardiac rehabilitation sports groups in a university [31] as well as with flyers and events at social centers and medical clinics [35].

Chronic Conditions of the Patients Included in the Studies of Mobile App Use

The apps used in this study included different conditions, such as asthma [29], heart failure [31,34], hypertension [30,33], and HIV [23,33]. The remaining apps did not focus on a specific illness [9,18,24,32,35]; however, the inclusion criteria for 2 of these studies included people suffering from multiple pathologies [9,32,33], and in another of the analyzed studies, the patients had to be recipients of solid organ transplants [24].

Designs Employed in the Studies

To compare results, 4 studies carried out randomized controlled trials [9,23,31,34]. The first of these compared treatment adherence in 2 groups: those who used the mobile app (experimental group) and those who did not (control group) [9]. The second study that conducted a randomized controlled trial compared treatment adherence between a control group that employed a mobile app with an experimental group that used an extended version of that same app [23]. The third randomized controlled study compared the interventions of 2 groups, mobile app and electronic pillbox, and under 2 conditions each: mobile app with a reminder, mobile app without a reminder, electronic pillbox with a reminder, and electronic pillbox without a reminder [34]. Finally, the fourth study compared the use of the app versus a paper diary [31].

XSL•FO
RenderX

Furthermore, 2 studies compared 2 independent samples. One of these compared the responses of persons older and younger than 55 years as its objective was to verify the differences between the effectiveness and ease of use of the app between these 2 groups [18]. Another compared the effectiveness of various mobile apps [35].

The remaining 5 studies [24,29,30,32,33] described assessment of the mobile apps by the patients

Time of Use of the Mobile Apps

The time that the participants used the mobile apps varied between 2 hours and 6 months, depending upon the study [33,35].

A description of these issues is in Table 1.

Table 1. Details of the included studies.

Authors and country	Objective	Participants	Chronic condition	Design (Duration)
Anglada-Martínez et al, Spain [33]	Evaluate 1 Web and smartphone-based medication self-management platform, named MedPlan.	N=42; average age: 56 years	Hypertension "and" or "or" dyslipidemia and HIV	Transversal (6 months)
Burbank et al, United States [29]	Examine the viability of a mobile application for adolescents with asthma.	N=20; adolescents; average age: 13.5 years	Asthma	Transversal (8 weeks)
Fallah and Yasini, France [18]	Design and evaluate a mobile medication reminder app.	N=60; <55 years: N=30; and >55 years: N=30	__[a]	Transversal (—)
Goldstein et al, United States [34]	Compare the adherence of 2 interventions, electronic pillbox and mobile apps, under experimental conditions with and without medication reminders, in addition to evaluating the viability and effectiveness of each.	N=58; elderly adults; average age: 69 years	Heart failure	Randomized controlled (28 days)
Grindrod et al, Canada [35]	Explore the ease of use and usefulness of existing mobile apps for handing medication in elderly adults.	N=35; >50 years; average age: 67 years	—	Transversal (2 hours)
Kang and Park, South Korea [30]	Develop a mobile application for managing hypertension and evaluate its usefulness, user satisfaction and adherence to medication.	N=38; average age: 56 years	Hypertension	Transversal (4 weeks)
Mertens et al, Germany [31]	Analyze if mobile application to support the therapy management will be accepted by elderly patients with chronic conditions and would improve their therapy adherence.	N=24; average age: 73.8 years	Coronary heart disease or myocardial infarction	Randomized controlled (84 days)
Mira et al, Spain [32]	Design, develop, and evaluate a mobile app that enables safer use of medication in elderly patients who take multiple medications.	N=61; elderly adults; average age: 68.8 years	Pluripathology	Transversal (—)
Mira et al, Spain [9]	Design, implement and evaluate a mobile app for self-management of medication in elderly patients who take multiple medications.	N=99; >65 years; experimental group: N=51; and control group: N=48	Pluripathology	Randomized controlled (3 months)
Perera et al, New Zealand [23]	Examine the effectiveness of a mobile application for facilitating treatment adherence to combined antiretroviral therapy.	N=28; average age: 46; experimental group: N=17; and control group: N=11	HIV	Randomized controlled (3 months)
Shellmer et al, United States [24]	Design a mobile application for improving treatment adherence in adolescent recipients of solid organ transplants and evaluate its acceptance, ease of use and satisfaction.	N=7; adolescents; + 9 caregivers	Recipients of solid organ transplants	Transversal (6 weeks)

[a]Missing data.

XSL•FO
RenderX

Functions of the Mobile Apps

The contents of the mobile apps included reminders for taking medication; some of these studies did so with alarms (visual and audio) that the patients had previously recorded [9,18,24,29-31,33]. When the alarm sounded, they had to confirm that they had taken the medication [9,23,24,29,31,33], and the apps notified their caregivers when the users failed to indicate that they had indeed taken the medication [9,24]. The apps provided instructions on how to take the medication [9,18,24,31-33], general information about the treatments and medication [18,30-33], education about the illness [24,29], and recommendations on healthy habits [9,30].

Some more specific functions of each app included reminders with alarms for doctors' appointments [30], blood pressure records [30], or their symptoms in general [29] and images of the medications taken to distinguish them when the time came for them to be taken, thus increasing patient safety [9,31,33]. The TUMEDICINA app (APPANDABOUT, SL) enabled scanning of the bar codes on medication containers to gain information about the intended therapeutic objectives, verbal instructions on how and when to take them, interactions with other medications, expiration dates, and storage indications. All this information was stored as audio recordings [32]. The app for HIV patients contained a 24-hour medication clock for the control and experimental groups. The latter used an extended version of the app that additionally included personalized images about the level of medication and the level of immunoprotection within the patient's body [23]. The Teen Pocket PATH app had one version for caregivers and another for patients and included general information such as telephone help numbers [24].

Profile of Mobile Apps Design Participants

In 6 of the studies, the design of the app was made from patient data compiled with qualitative techniques, such as in nominal groups [9,18,24,29,30,32]. In addition, health professionals participated in the app design in 3 of them [9,18,29], and technology specialists also participated in 1 [19]. In another study, in addition to including participation from the patients who were subsequently going to use the app, the design also kept their caregivers in mind [24]. In another study, the app was designed exclusively by technology experts [30].

Mobile App Availability

In 5 of the 11 studies, the mobile apps were available in both Android and iOS versions [9,29,32,34,35], whereas 4 were only available for Android [18,23,24,30,33] and 2 were only available for iOS [31]. Furthermore, among the studies, 1 study compared mobile apps of Android and iOS environments [35]. In 3 studies, the participants downloaded the app on their mobile phones [23,29,30] and in 4 studies, the users were offered either iPads or tablets with the downloaded app [9,24,31,35].

Reference Measures for Evaluating Treatment Adherence

The questionnaires administered for evaluating treatment adherence were the Modified Morisky Scale [30]; the Morisky Medication Adherence Scale along with a questionnaire for evaluating the rates of lost doses and medication errors [9]; Medication Adherence Report Scale [23]; Simplified Medication Adherence Questionnaire [33]; the subjective adherence measure A14 scale [31]; the Asthma Control Test for evaluating the impact of asthma on daily functions, frequency of shortness of breath, frequency of asthmatic symptoms at night, frequency of using rescue medicines, and general control of asthma; and the Child Asthma Self-Efficacy for determining the prevention and management of asthma attacks [29]. The other methods used included pharmacy dispensers, measuring the quantity of virus in the blood plasma of each HIV patient [23], and dividing the number of medications that were marked as having been taken by the number of medicines prescribed [34].

Quantitative Measures for Evaluating Mobile App Functions

The questions used for evaluating the mobile apps included the Post Study System Usability Questionnaire [24] and the System Usability Scale, which determined the use of the app [35]. One study administered a questionnaire to evaluate the app's effectiveness and ease of use in which the questions for evaluating the ease were extracted from the System Usability Scale [18]; 1 study created a questionnaire to evaluate the acceptance, usefulness, satisfaction, willingness to recommend the app to other persons, and the opinion the users held about it [34]; and 1 study evaluated the usefulness by using the questionnaire on perceived usefulness by Davis and the satisfaction by means of a questionnaire that evaluated satisfaction with each of the app's contents [30]. Another study evaluated the use of the app according to the number of times that each participant examined each of the contents in the app and the amount of time invested in each content. Furthermore, a questionnaire was administered with questions about the app's satisfaction, perceived utility, ease of use, visual appeal, and discretion and about the information it provided [23]. In another study, a questionnaire was administered that evaluated the app's characteristics and operation [32]. Finally, 1 study assessed usability and satisfaction through self-reported questionnaires [33].

Qualitative Measures for Evaluating Mobile App Functions

Overall, 7 studies compiled patient data using qualitative techniques wherein questions were asked about the satisfaction, usefulness, ease of use, acceptance and the contents of the apps [9,24,29,31,32,35].

Mobile App Effectiveness in Treatment Adherence

Furthermore, 7 studies confirmed that the mobile app increased treatment adherence [9,23,29-31,33,34], and in 5 of them, the differences in adherence before and after the study were statistically significant [9,23,30,31,33]. The study that compared the intervention of the mobile app with that of the electronic pillbox did not find significant differences between the type of device used or the reminders and treatment adherence; however, the participants declared that they preferred the mobile device [34]. Another study did not find statistically significant differences in the control of asthma before and after the study, although the patients with uncontrolled asthma before the study did show a significant increase in their scores. Mean scores on asthma self-efficacy before and after the study increased but

XSL•FO
RenderX

were not significant. However, there was a significant increase in preventing an asthma attack [29]. In addition, 3 studies found that the mobile app reduced the occurrence of missed dose significantly [9,31,33]. In addition, the device decreased medication mistakes only in people who had reported committing 2 or more errors before the study [9].

Satisfaction With the Mobile Apps

The participants declared that that they were satisfied with the app in all 7 of the studies that included this measure [9,23,24,29,30,32,33]. They were more satisfied with the functions that helped them to promote treatment adherence such as reminders and recording symptoms and medication information [30,35]. People who rated the highest were those who organized their medication in pillboxes, took notes on medication containers, and took less than six every day [32]. Moreover, experimental groups who used mobile apps were more satisfied compared with control groups with other devices [23,34].

Other Evaluated Elements

Ease of use was estimated in 6 studies [9,18,23,24,32,35], and in 4 of them, the participants stated that the app was easy to use [9,18,24,32]. Furthermore, 1 study confirmed that there were no statistically significant differences in the ease of use between those younger and older than 50 years [18], whereas another found no statistically significant differences between persons who used mobile phones or browsed the internet with those who did not [32]. In 1 study that compared various apps, only 1 of the apps received scores for ease of use that were lower than the remaining apps. Moreover, people rated the experience of using the mobile apps as difficult, although that changed when they learned how to use them [35].

In 5 studies, the participants stated that these mobile apps were useful [23,24,30,32,34]. In addition, in 1 study, the participants suggested that the app would be even more useful if it added the option of an alarm as a reminder for taking medication [23]. Yet another study demonstrated that the ideal app would be one that helped foster treatment adherence and, furthermore, provided information about the illness and its treatment [35].

Finally, 1 study [23] that compared a reduced version of an app (control group) with an extended version (experimental group) found that the participants from the experimental group rated their app as more informative, more visually appealing, and more of a motivator for promoting adherence to treatment in comparison with the control group, and almost all the participants would recommend the mobile app to their friends.

A description of these issues is in Table 2.

XSL•FO
RenderX

Table 2. Details of the apps used in the included studies.

Study	App functions and design	Medication adherence measure	Measure for evaluating app	App evaluation
Anglada-Martínez et al [33]	MEDPLAN. Drugs information, medication reminder alarm system, where patients confirm whether they have taken the drug or not. App designed by health professionals.	Simplified Medication Adherence Questionnaire (SMAQ), pharmacy refill method and number of days with missing dose.	Usability and satisfaction assessed through self-reported questionnaires.	When adherence was measured using the SMAQ, treatment adherence improved during the intervention phase (19.4%; $P<.05$), and the number of days with missed doses decreased ($P<.05$). The mean satisfaction score for Medplan was 7.2 ± 2.7 out of maximum of 10 points. 71.4% of participants said they would recommend the App to a friend, and 88.1% wanted to continue using it. They thought the application could be more useful in patients on polypharmacy, at the beginning of a treatment, for caregivers or for the elderly population.
Burbank et al [29]	Medication reminder, reminder for recording symptoms, feedback on its adherence and education about asthma. The App was designed by patients and health professionals.	Asthma Control Test. Child Asthma Self-Efficacy Questionnaire.	Questions about satisfaction.	In spite of the improvement in the control of asthma before and after the study, there were no significant differences ($P=.53$). However, the scores improved significantly for those who did not control asthma before the intervention ($P=.03$). Mean scores on self-efficacy before and after the study increased, but were not significant ($P=.13$). Although there were significant differences in preventing an asthma attack ($P=.04$). Satisfaction: 93%
Fallah and Yasini [18]	Reminders via alarms, instructions and information about medication. The App was designed by patients, health professionals, and technology specialists.	—[a]	Questionnaire for evaluating the application's effectiveness and ease of use. The questions for evaluating its ease of use were taken from the System Usability Scale adapted for mobile applications.	No significant differences were found between the effectiveness or ease of use in either age group (greater and younger than 50). Both groups found the app effective and easy to use.
Goldstein et al [34]	—	Electronic pillbox: opening the pillboxes. Mobile application: electronic self-reports. The number of medications taken was divided by the number of medications prescribed.	Questionnaire for evaluating the acceptance, usefulness, satisfaction, willingness to recommend it and their opinion about the device.	Improves treatment adherence with both interventions (80%). No significant differences were found between the type of device and adherence ($P=.87$), neither were there between the condition and adherence ($P=.48$). Those in the mobile application group awarded higher scores on acceptance and usefulness of their device ($P<.001$). All participants preferred the intervention of the mobile application.
Grindrod, Li and Gates [35]	—	—	System Usability Scale. Questions in a group session: ease of use, user experiences, expected adoption, concerns about the potential for data entry errors, perceived quality of the provided information and preferences for the different characteristics.	The Pocket Pharmacist application received an ease of use score that was significantly lower when compared to the remaining applications ($P<.001$). They initially rated the experience of using the applications as frustrating, although that changed when they learned how to use them. They would use these applications if they someday needed to due to cognitive or health problems. The ideal application would possess characteristics that helped foster adherence and provide information.

XSL•FO
RenderX

Study	App functions and design	Medication adherence measure	Measure for evaluating app	App evaluation
Kang and Park [30]	HYPERTENSION MANAGEMENT APP. Reminders with alarms for taking medication and doctor's appointments, recording blood pressure, recommendations about lifestyle and information on medication. The App was designed by patients and experts.	Modified Morisky Scale.	Questionnaire with a scale from 1 to 5 that evaluated perceived usefulness and satisfaction with each of the application's contents.	The average scores on adherence increased significantly before and after the study from 4.2 to 5.2 out of a maximum of 6 points (P=.001). Perceived usefulness: 3.7. Satisfaction: 3.8 for medication reminders, 3.2 for alarms, 4.3 for recording blood pressure, 3.1 for the information sent, 3.4 for recommendations, and 3.8 for education about medication.
Mertens et al [31]	MEDICATION PLAN. Reminders via alarms, instructions and information about medication. The App was designed by health professionals.	Subjective adherence was determined by the A14-scale. Objective adherence was measured by number of medications each participant had to take each day.	Semistructured interviews.	The mean for subjectively assessed adherence there was a significant increase after the interventional phase from 50 to 54 out of a maximum of 56 points (P<.001). The app showed significant adherence for medication intake (P=.03). The majority of participants (n=22) stated that they would like to use the medication app in their daily lives.
Mira et al [32]	TUMEDICINA. Scans the bar codes on the medication box to provide information about its therapeutic objective, indications for taking it, interactions with other medications and its date of expiration. This information is stored as audio recordings. The App was designed by patients.	—	Group session and individual questionnaire for evaluating the characteristics and operation of the application.	The characteristics rated highest were the simplicity and clarity of the verbal messages (96.7%), the usefulness of the verbal messages (93.4%) and the clarity of the information provided (95.1%). No significant differences were found in the assessment of the satisfaction between patients with or without experience of using mobile telephones or browsing the Internet (P=.88). The people who rated the application the highest were persons who organized their medication in pillboxes, took notes on medication containers and took less than six drugs every day. Satisfaction: 8.3 out of 10.
Mira et al [9]	ALICE. Reminders with alarms for taking medication and carrying out healthy habits, images of drugs, instructions on how to take medication, SMS sent to caregivers in cases where the medication is not taken. The App was designed by patients, health professionals, and technology specialists.	Morisky Medication Adherence Scale. Questionnaire for evaluating rates of missed doses and medication errors.	Questions for evaluating the application: satisfaction, ease of use, performance, usefulness, reliability, acceptance, design, simplicity, accessibility, if they would recommend it and if it afforded them independence.	Treatment adherence improved in the experimental group (28%; P<.001) and in a lower rate of omitted doses (27.3%; P=.02). The application was not effective in reducing the rate of medication errors, it only decreased in patients who had reported committing 2 or more errors before the study (41.2%). Satisfaction: 8.5 out of 10. Persons without experience of information technologies said that using the application was not complicated.
Perera et al [23]	The application used by the control group contained a 24-hour medication watch. For the experimental group, in addition to the watch, it contained personalized messages about the levels of medication and immunoprotection in the patient's body.	Medication Adherence Report Scale. Pharmacy prescriptions filled. HIV viral load.	Questionnaire for evaluating the satisfaction, perceived usefulness, ease of use, visual appeal, discretion and provision of information.	Greater treatment adherence in the experimental group according to the scores on the Medication Adherence Report Scale (40%; P=.03) and according to the viral load HIV (19%; P=.02). However, there were no significant differences in the pharmacy dispensing data (P=.18). The experimental group participants were more satisfied with the application than the control group and they rated it as informative, attractive and motivating. 79% said that adding the option of an alarm to remind about taking medication would be useful. 81% of the experimental group would recommend the application.

XSL•FO
RenderX

Study	App functions and design	Medication adherence measure	Measure for evaluating app	App evaluation
Shellmer et al [24]	TEEN POCKET PATH. Reminder of what medication must be taken and in what dose, confirmation that it had been taken, information about the type of transplant received and general information, such as telephone help lines. Caregivers received information as to whether the adolescents had taken their medication. The App was designed by patients and caregivers.	—	Post Study System Usability Questionnaire. Questions during one session: ease of use, viability, satisfaction, usefulness, simplicity of the reminder, warning messages sent to the caregivers and perceptions about long-term use of the application.	Users and caregivers found the application easy to use, effective, useful and they were satisfied with it. The caregivers said that they felt less need to constantly ask the adolescents about whether or not they had taken their medication.

[a]Missing data.

Discussion

Principal Findings

These results indicate that mobile apps help promote treatment adherence [9,23,29,30,34]. However, when considering the sample size and time of use of the mobile apps under natural conditions, new studies with longer use times than the apps merit consideration to find out whether an accommodation effect exists that has a negative effect upon adherence, for example, after more than 12 or 18 months of using these apps [21,22,25].

One thing to keep in mind is that these studies focused exclusively on the lack of adherence caused involuntarily by the patient. They did not control participant variables of the locus control type or confidence or relationship with health professionals. Users of these apps who voluntarily and consciously rule out following the treatment can use these devices to gain greater credibility with their caregivers or health professionals by indicating in the app that their medication has been taken even when this is not the case. This is the same problem with traditional pillboxes and in research on therapeutic adherence [8].

The gold standard used for determining therapeutic effectives has been based on the use of reports by patients obtained using validated scales and widely used in research on adherence [9,23,29,30]. Only 1 study used a more objective and reliable measure of adherence, that of blood determinations [23].

The majority of patients stated that the mobile apps they had used were easy to use [9,18,24,32] and useful [23,24,30,32,34] and that, additionally, they were satisfied with their ease of use, navigation, and features [9,23,24,29,30,32]. The studies analyzed show that persons aged over 60 years do not encounter difficulties when using these apps and that, therefore, there are no barriers because of age [9,18,32]. In these cases, it should be pointed out that the apps had been designed with the intrinsic characteristics of the target population in mind, such as letter or image sizes [9,18,32]. In addition, it is important to consider that personal characteristics, such as computer literacy, health literacy, mental health status, and cultural background, are related with the use of mHealth apps [36].

It should be noted that in most of the studies, the mobile apps were designed especially for future users [9,18,24,29,30,32]. This indicates that the app design is made according to the needs of patients and has probably contributed not only to their effectiveness but also, above all, to satisfaction with the app.

The main contents in the apps to foster treatment adherence were reminders with alarms for taking the medication [9,18,30], information about the medication [18,30,32], and medication-tracking histories [9,23,25,29].

Although the level of knowledge about the illnesses or their treatments was not controlled in the studies carried out, one could expect that using these apps contributes to greater knowledge about the disease and the drugs that are taken every day. In some cases, these apps include information about drug storage and about potential (the most frequent) drug interactions with other active ingredients or natural products [18,30,32]. This is a relevant aspect because the studies point out that the knowledge patients possess about their medication could be improved, and it should be an objective when these types of apps are designed. Education about treatment is especially important for those who commit more errors in its administration, such as those who use devices such as glucometers or inhalers, and for caregivers of minors [7,8].

Other app functions to promote adhering to the therapeutic regimen were reminders about leading a healthy lifestyle [9,30] or reminders about keeping appointments with physicians [30]. These functions, positively valued by patients, provide added value compared with traditional pillboxes.

This review shows that mobile apps are effective in promoting treatment adherence and that they contribute to patient safety by avoiding errors in the administration of their treatments. Owing to this, health professionals, such as physicians or pharmacists [37], should promote their use by recommending that their patients download them and then monitor how these apps are used, because simply downloading them does not ensure their full use [8,9].

Limitations

Among the possible limitations of this study, it should be mentioned that despite having carried out the search in the most important databases on medicine, it is probable that other

XSL•FO
RenderX

databases were not considered. In addition, although we used a wide range of descriptors to obtain a more precise strategy, there might be a specific keyword from a concrete area that was not controlled.

Furthermore, we did not include articles in languages other than English and Spanish nor did we consider abstracts from conferences.

Another limitation to highlight is the difficulty in compiling the results because of the wide heterogeneity of methodologies and results from the articles that were found.

This study evaluates the effectiveness of mobile apps as a method for overcoming errors by patients in managing medication. However, these mobile apps do not offer alternatives for controlling voluntary nonadherence by patients.

Comparison With Other Studies

We know that 1 in 5 elderly patients forget to take their medication or make mistakes when doing so [38]. The use of new technologies is a relevant method for overcoming the problems of lack of adherence to treatments, which results in harmful consequences for the health of patients and for those who are elderly. The effectiveness of the mobile apps could be because of the effects that alarms have on forgetfulness, as this is one of the main contributors to the lack of treatment adherence [25], but these apps must also be employed with patient safety in mind, for example, with information on how to avoid drug interactions, with information on how to properly store the medication, or with instructions on which foods the medication can and cannot be taken with.

Most of the studies focused on specific diseases, but all of them had a common approach toward chronic diseases [9,23,24,29,30,32,34]. Previous research found that lack of treatment adherence is more frequent in persons with chronic diseases because of the complexity of therapeutic regimens, regardless of age [5,9,32,36]. For this reason, solutions to lack of treatment adherence caused involuntarily by the patient must be personalized by considering the patient's profile and the posology, which have a more direct impact on the difficulties of taking medication.

The relevance of using smartphones to foster treatment adherence is also because of their acceptance, ease of use, and affordability [21,26]. These findings justify that elderly people, when the app has been designed with their needs in mind, are not a barrier as some of the reviewed studies suggest.

Park et al [21] found that positive and personalized feedback resulted in positive effects on medication adherence. This is the function that digital pillboxes perform. Personalizing alarms could contribute to their effectiveness and to that effectiveness lasting for longer periods.

Other studies have evaluated the effectiveness of other technological methods by which treatment adherence can be enhanced, such as telephone calls or SMS [20,21,22]. Although these are just as effective, they involve high costs [20,26]. Furthermore, these interventions only take into account reminders for taking the medication, whereas mobile apps provide more content, such as educational interventions [22].

In addition to mobile apps found for fostering treatment adherence, there are also apps for promoting adherence to other therapeutic regimens, with reminders for leading an appropriate lifestyle, reminders for keeping doctors' appointments, and monitoring other health information (eg, supplements and manage pets), among others [26,38]. The integration of these functions should be considered when designing new apps for virtual pillboxes.

Future Research

From these results, recommendations for the design of future apps can also be deduced when considering the contents valued highest by the patients. Park et al [37] have found that features appreciated by users are app performance and practical aspects, helpful reminders and notifications, monitoring other health information, versatility of medication information input and display, and supporting health care visits [37]. Standing out among these are the flexible management of alarms that warn about taking medication and education about the type, use of, and precautions about the medication that they take [23,35]. Conversely, these functions have better value for the participants when they use simple interfaces. For this reason, mobile apps are easy to use and people make more use of them [35,39]. Furthermore, it should be emphasized that the future users of the apps must participate in their design to focus on their necessities [9,18,24,29,35,39].

The majority studies included in this review evaluated treatment adherence by validated scales such as the Morisky Medication Adherence Scale [9,30]. Future studies should incorporate objective measures, for example, the most common measure is blood test [23].

In addition, patient safety should be considered in these mobile apps because these help them to manage their medication and they could make mistakes when taking their drugs [9].

Finally, the studies with longer use times of the apps are considered necessary to integrate the mobile apps in their daily routine and examine their effectiveness for treatment adherence in the long term [17-19,22].

Conclusions

Mobile apps prevent forgetting about medication and incorrect administration and, thus, contribute to patient safety. In the future, these apps should include personalization of the personal conditions and posology of the medication the patient takes.

Acknowledgments

This study was financed by the Conselleria de Educación, Generalitat Valenciana Excellence Groups Prometeo/2017/173.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Levels of evidence and degrees of recommendation.

[PDF File (Adobe PDF File), 156KB - jmir_v21i6e12505_app1.pdf]

Multimedia Appendix 2

Assessment of the internal quality of the design of the studies.

[PDF File (Adobe PDF File), 138KB - jmir_v21i6e12505_app2.pdf]

References

1. Sabaté E. Adherence to Long-term Therapies: Evidence for Action. Geneva: World Health Organization; 2003.
2. Dilla T, Valladares A, Lizán L, Sacristán JA. [Treatment adherence and persistence: causes, consequences and improvement strategies]. Aten Primaria 2009 Jun;41(6):342-348 [FREE Full text] [doi: 10.1016/j.aprim.2008.09.031] [Medline: 19427071]
3. García P. [University of Granada Faculty of Educational Sciences]. 2008. [Patient's knowledge about their medications] URL: https://hera.ugr.es/tesisugr/1771557x.pdf
4. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: a systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 2;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
5. Aguirrezabala J, Aguado M, Aizpurua I, Albizuri M, Alfonso I, Armendáriz M. [Electronic headquarters - Basque Government]. 2011. [Adherence to pharmacological treatment in chronic pathologies] URL: http://www.euskadi.eus/contenidos/informacion/cevime_infac_2011/es_def/adjuntos/infac_v19_n1.pdf [WebCite Cache ID 70SOnuX4z]
6. Bosworth HB, Zullig LL, Mendys P, Ho M, Trygstad T, Granger C, et al. Health information technology: meaningful use and next steps to improving electronic facilitation of medication adherence. JMIR Med Inform 2016;4(1):e9 [FREE Full text] [doi: 10.2196/medinform.4326] [Medline: 26980270]
7. Pérez-Jover V, Mira JJ, Carratala-Munuera C, Gil-Guillen VF, Basora J, López-Pineda A, et al. Inappropriate use of medication by elderly, polymedicated or multipathological patients with chronic diseases. Int J Environ Res Public Health 2018 Feb 10;15(2):310 [FREE Full text] [doi: 10.3390/ijerph15020310] [Medline: 29439425]
8. Mira JJ, Lorenzo S, Guilabert M, Navarro I, Pérez-Jover V. A systematic review of patient medication error on self-administering medication at home. Expert Opin Drug Saf 2015 Jun;14(6):815-838. [doi: 10.1517/14740338.2015.1026326] [Medline: 25774444]
9. Mira JJ, Navarro I, Botella F, Borrás F, Nuño-Solinís R, Orozco D, et al. A Spanish pillbox app for elderly patients taking multiple medications: randomized controlled trial. J Med Internet Res 2014;16(4):e99 [FREE Full text] [doi: 10.2196/jmir.3269] [Medline: 24705022]
10. Chan HK, Hassali MA, Lim CJ, Saleem F, Tan WL. Using pictograms to assist caregivers in liquid medication administration: a systematic review. J Clin Pharm Ther 2015 Jun;40(3):266-272. [doi: 10.1111/jcpt.12272] [Medline: 25865563]
11. Mira JJ, Orozco-Beltrán D, Pérez-Jover V, Martínez-Jimeno L, Gil-Guillén VF, Carratala-Munuera C, et al. Physician patient communication failure facilitates medication errors in older polymedicated patients with multiple comorbidities. Fam Pract 2013 Feb;30(1):56-63. [doi: 10.1093/fampra/cms046] [Medline: 22904014]
12. Schwartz JK. Pillbox use, satisfaction, and effectiveness among persons with chronic health conditions. Assist Technol 2017;29(4):181-187. [doi: 10.1080/10400435.2016.1219884] [Medline: 27689861]
13. e Souza FR, da Silva Santana C. A descriptive study about the use of pillboxes by older adults. Health 2013;05(12):103-109. [doi: 10.4236/health.2013.512A014]
14. Sanders MJ, van Oss T. Using daily routines to promote medication adherence in older adults. Am J Occup Ther 2013;67(1):91-99. [doi: 10.5014/ajot.2013.005033] [Medline: 23245787]
15. Zeller A, Schroeder K, Peters TJ. Electronic pillboxes (MEMS) to assess the relationship between medication adherence and blood pressure control in primary care. Scand J Prim Health Care 2007 Dec;25(4):202-207 [FREE Full text] [doi: 10.1080/02813430701651954] [Medline: 17924286]
16. Haase J, Farris KB, Dorsch MP. Mobile applications to improve medication adherence. Telemed J E Health 2016 Jun 1;23(2):75-79. [doi: 10.1089/tmj.2015.0227] [Medline: 27248315]
17. Hayes TL, Hunt JM, Adami A, Kaye JA. An electronic pillbox for continuous monitoring of medication adherence. Conf Proc IEEE Eng Med Biol Soc 2006;1:6400-6403 [FREE Full text] [doi: 10.1109/IEMBS.2006.260367] [Medline: 17946369]
18. Fallah M, Yasini M. A medication reminder mobile app: does it work for different age ranges? Stud Health Technol Inform 2017;235:68-72. [Medline: 28423757]

XSL•FO
RenderX

19. Rivero F. Asian Media Information and Communication Centre. 2017. [Mobile Report in Spain and the World 2017] URL: https://www.amic.media/media/files/file_352_1289.pdf
20. Belzer ME, Kolmodin MK, Clark LF, Huang J, Olson J, Kahana SY, Adolescent Medicine Trials Network for HIV/AIDS Interventions. Acceptability and feasibility of a cell phone support intervention for youth living with HIV with nonadherence to antiretroviral therapy. AIDS Patient Care STDS 2015 Jun;29(6):338-345 [FREE Full text] [doi: 10.1089/apc.2014.0282] [Medline: 25928772]
21. Park LG, Howie-Esquivel J, Dracup K. A quantitative systematic review of the efficacy of mobile phone interventions to improve medication adherence. J Adv Nurs 2014 Sep;70(9):1932-1953. [doi: 10.1111/jan.12400] [Medline: 24689978]
22. Tao D, Xie L, Wang T, Wang T. A meta-analysis of the use of electronic reminders for patient adherence to medication in chronic disease care. J Telemed Telecare 2015 Jan;21(1):3-13. [doi: 10.1177/1357633X14541041] [Medline: 25147178]
23. Perera AI, Thomas MG, Moore JO, Faasse K, Petrie KJ. Effect of a smartphone application incorporating personalized health-related imagery on adherence to antiretroviral therapy: a randomized clinical trial. AIDS Patient Care STDS 2014 Nov;28(11):579-586 [FREE Full text] [doi: 10.1089/apc.2014.0156] [Medline: 25290556]
24. Shellmer DA, Dew MA, Mazariegos G, DeVito Dabbs A. Development and field testing of Teen Pocket PATH(®), a mobile health application to improve medication adherence in adolescent solid organ recipients. Pediatr Transplant 2016 Feb;20(1):130-140 [FREE Full text] [doi: 10.1111/petr.12639] [Medline: 26916967]
25. Simón E, Vañó S. [Health: Trends and Applications]. Madrid, España: Fundación Gaspar Casal; 2016.
26. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behaviour change? An evidence review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
27. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
28. Scottish Intercollegiate Guidelines Network. SIGN 50: A Guideline Developers Handbook. Edinburgh: NHS Quality Improvement Scotland (NHS QIS); 2008.
29. Burbank AJ, Lewis SD, Hewes M, Schellhase DE, Rettiganti M, Hall-Barrow J, et al. Mobile-based asthma action plans for adolescents. J Asthma 2015;52(6):583-586 [FREE Full text] [doi: 10.3109/02770903.2014.995307] [Medline: 25494553]
30. Kang H, Park H. A mobile app for hypertension management based on clinical practice guidelines: development and deployment. JMIR Mhealth Uhealth 2016 Feb 2;4(1):e12 [FREE Full text] [doi: 10.2196/mhealth.4966] [Medline: 26839283]
31. Mertens A, Brandl C, Miron-Shatz T, Schlick C, Neumann T, Kribben A, et al. A mobile application improves therapy-adherence rates in elderly patients undergoing rehabilitation: A crossover design study comparing documentation via iPad with paper-based control. Medicine (Baltimore) 2016 Sep;95(36):e4446 [FREE Full text] [doi: 10.1097/MD.0000000000004446] [Medline: 27603339]
32. Mira JJ, Guilabert M, Carrillo I, Fernández C, Vicente MA, Orozco-Beltrán D, et al. Use of QR and EAN-13 codes by older patients taking multiple medications for a safer use of medication. Int J Med Inform 2015 Jun;84(6):406-412. [doi: 10.1016/j.ijmedinf.2015.02.001] [Medline: 25731975]
33. Anglada-Martínez H, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Sequeira E, Aragunde V, et al. Feasibility and premilinary outcomes of a web and Smartphone-based medication self-management platform for chronically ill patients. J Med Syst 2016 Apr;40(4):99. [doi: 10.1007/s10916-016-0456-y] [Medline: 26872781]
34. Goldstein CM, Gathright EC, Dolansky MA, Gunstad J, Sterns A, Redle JD, et al. Randomized controlled feasibility trial of two telemedicine medication reminder systems for older adults with heart failure. J Telemed Telecare 2014 Sep;20(6):293-299. [doi: 10.1177/1357633X14541039] [Medline: 24958355]
35. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
36. Wildenbos GA, Peute LW, Jaspers MW. Influence of human factor issues on patient-centered mHealth apps' impact; where do we stand? Stud Health Technol Inform 2016;228:190-194. [Medline: 27577369]
37. Park JYE, Li J, Howren A, Tsao NW, De Vera M. Mobile phone apps targeting medication adherence: quality assessment and content analysis of user reviews. JMIR Mhealth Uhealth 2019 Jan 31;7(1):e11919 [FREE Full text] [doi: 10.2196/11919] [Medline: 30702435]
38. Mira JJ, Navarro IM, Guilabert M, Aranaz J. [Frequency of medication errors by patients]. Rev Panam Salud Publica 2012 Feb;31(2):95-101 [FREE Full text] [Medline: 22522870]
39. Carrillo I, Guilabert M, Pérez-Jover V, Mira JJ. [Assessment of two applications of medication self-management in older patients. Qualitative study]. Rev Calid Asist 2015;30(3):142-149. [doi: 10.1016/j.cali.2015.02.005] [Medline: 25843349]

Abbreviations

CRT: controlled and randomized trial
mHealth: mobile health
MEMS: Medication Event Monitoring System
SMS: short message service

XSL•FO
RenderX

Edited by CL Parra-Calderón; submitted 15.10.18; peer-reviewed by C Alvarez, B Mason, J Farzi; comments to author 26.01.19; revised version received 28.03.19; accepted 11.05.19; published 18.06.19.

Please cite as:
Pérez-Jover V, Sala-González M, Guilabert M, Mira JJ
Mobile Apps for Increasing Treatment Adherence: Systematic Review
J Med Internet Res 2019;21(6):e12505
URL: http://www.jmir.org/2019/6/e12505/
doi:10.2196/12505
PMID:31215517

©Virtudes Pérez-Jover, Marina Sala-González, Mercedes Guilabert, José Joaquín Mira. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Tobacco Use Behaviors, Attitudes, and Demographic Characteristics of Tobacco Opinion Leaders and Their Followers: Twitter Analysis

Kar-Hai Chu[1*], PhD; Anuja Majmundar[2*], MBA, MA; Jon-Patrick Allem[2*], PhD; Daniel W Soto[2*], MPH; Tess Boley Cruz[2], PhD; Jennifer B Unger[2*], PhD

[1]University of Pittsburgh, Pittsburgh, PA, United States

[2]University of Southern California, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
Kar-Hai Chu, PhD
University of Pittsburgh
230 McKee Place, Suite 600
Pittsburgh, PA,
United States
Phone: 1 412 692 2578
Email: chuk@pitt.edu

Abstract

Background: Tobacco-related content on social media is generated and propagated by opinion leaders on the Web who disseminate messages to others in their network, including *followers*, who then continue to spread the information. Opinion leaders can exert powerful influences on their followers' knowledge, attitudes, and behaviors; yet, little is known about the demographic characteristics and tobacco use behavior of tobacco opinion leaders on the Web and their followers, compared with general Twitter users.

Objective: In this study, we hypothesized that opinion leaders use more tobacco products and have higher nicotine dependence than the other 2 groups (eg, followers and general Twitter users) and that followers—those who spread messages by opinion leaders—would more likely be in demographic groups that are vulnerable to tobacco marketing influence (eg, young adults and lower educational attainment).

Methods: We constructed the social networks of people who tweet about tobacco and categorized them using a combination of social network and Twitter metrics. To understand the characteristics of tobacco opinion leaders and their followers, we conducted a survey of tobacco opinion leaders, their followers, and general Twitter users. The sample included 347 opinion leaders, 567 followers, and 519 general users. The opinion leaders had a median of 1000 followers, whereas followers and general users had fewer than 600 followers.

Results: Opinion leaders were more likely than their followers to report past month use of tobacco products; followers, in turn, were more likely to report past month use of these products than general Twitter users. The followers appeared to be an especially vulnerable group; they tended to be younger (mean age 22.4 years) and have lower education compared with the opinion leaders and general users.

Conclusions: Followers of Twitter tobacco opinion leaders are a vulnerable group that might benefit from antitobacco education to counter the protobacco communications they see on social media.

(J Med Internet Res 2019;21(6):e12676) doi:10.2196/12676

KEYWORDS

tobacco; social media; online social networking; peer influence; social networking

XSL•FO
RenderX

Introduction

Social media has emerged as a popular forum for tobacco product users and prospective new users to learn about and discuss nicotine and tobacco products and for businesses to promote these products. Previously identified themes of tobacco-related social media posts include marketing and promotions by manufacturers, posts by tobacco users about their own tobacco experiences, discussions about flavors, and debates about tobacco product regulations [1-3]. Tobacco-related social media posts also contain content that could entice youth and nontobacco users to try tobacco products, including cartoons and other youth-oriented themes [4,5], hookah paired with alcohol in social settings [6], little cigar or cigarillos paired with marijuana [7], and pictures of people blowing large clouds of e-cigarette aerosol [8,9].

As social media messages about tobacco products can influence perceptions about the health effects and potential harms of these products and the social norms of use [10], it is important to understand as much as possible about these messages and the people who are disseminating them. Although previous research has focused on the content of tobacco-related messages on social media, less is known about the people who generate, receive, and propagate those messages. Previous analyses of information flow in Twitter have shown that a small number of *elite* users (approximately 20,000 people) generate nearly half of the tweets [11]. Only about 15.02% (1377/9165) of tweets received by ordinary users are directly from traditional mass media sources (eg, Cable News Network); most are filtered through opinion leaders who selectively retweet information from mass media sources [11]. Twitter opinion leaders—people who occupy central positions in their social networks, have numerous well-connected followers, have social status and credibility, and are emulated by their followers—are important members of Web-based communities as they control which information diffuses through social networks [12].

Opinion leaders discussing tobacco products on Twitter can potentially influence their followers to try new products, adopt beliefs about the relative harm of tobacco products, and support or oppose tobacco control policies. Thus, it is important to understand who these tobacco-related opinion leaders are and how their personal attitudes and behaviors might be influencing the discourse on the Web about tobacco products.

Intervention or education programs can benefit by leveraging opinion leaders to champion their ideas [13].

This study identified tobacco-related opinion leaders on Twitter by combining Twitter user metadata with techniques in social network analysis to develop a more comprehensive definition of opinion leaders. We then conducted surveys of these opinion leaders, people who follow these opinion leaders, and general Twitter users who are not engaged in tobacco-related discussions. We compared social network characteristics, demographic characteristics, tobacco product use, and nicotine dependence to identify differences among opinion leaders, followers, and general Twitter users. Opinion leaders, in this research, are operationalized as individuals strategically situated in their social networks. Their messages are disseminated widely via *shares* or *retweets* as they are viewed as subject matter experts. Research suggests that such high involvement and engagement in specific topics leads individuals to raise awareness about those topics and transition to polytobacco product use [14,15]. In keeping with this behavior, we hypothesized that opinion leaders would use more tobacco products and have higher nicotine dependence than the other 2 groups. We also hypothesized that followers, operationalized as individuals who predominantly follow and disseminate social media messages of opinion leaders, would more likely be in demographic groups that are vulnerable to the influence of tobacco marketing (eg, young adults and racial or ethnic minorities).

Methods

Data Collection

Twitter data were obtained with a custom Java 7 program based on Twitter4J v.4.0.3 that continuously accesses the Twitter streaming application programming interface (API) v.1.1 and collects tweets that contain any 1 of over 200 tobacco-related keywords, for example, cigarette, e-cigarette, or vape (see Multimedia Appendix 1). Twitter data were collected from March 2015 to March 2016. Along with the text of the tweet, the data include Twitter metadata such as the user name of the person who posted the tweet and whether the tweet was an original tweet or a retweet. This information was used to construct the retweet network by retrieving the data of every user who posted tobacco-related content.

Solely using Twitter metrics to identify opinion leaders can be misleading. Nontraditional accounts (eg, celebrities) can distort actual user classifications, whereas the number of followers is more likely to measure popularity rather than influence [16,17]. Therefore, we applied a combination of social network analysis–clustering algorithms and Twitter metrics to classify 3 types of individuals as follows: an opinion leader, a follower, and a general Twitter user. Opinion leaders and followers are additionally defined as users in our data who had posted tobacco-related content compared with general Twitter users who did not. First, a network was generated by linking users who had retweeted another user; this resulted in a retweet network defined by ties between the person who posted a tweet and the person who retweeted it. From this network, clusters were identified by conducting a modularity analysis. Modularity helps identify clusters within a network by grouping nodes (ie, Twitter users) that have more connections (ie, retweets) with others within a group than those outside of the group [18]. After the clusters were identified, opinion leaders were chosen as those who had been retweeted the most; followers were identified within each cluster as those who had retweeted others the most. Independently, general Twitter users were found by the Twitter API's get-user-status function, which returns users who have recently posted a tweet about any topic (not just tobacco).

This method produced a convenience sample of 347 opinion leaders, 567 followers, and 519 general users. We sent Twitter private messages to potential participants inviting them to complete the survey. Each private message contained a unique,

XSL•FO
RenderX

randomly generated link to a RedCap site where the survey was hosted. Clicking on the link identified the respondent as an opinion leader, follower, or general user who had been invited to complete the survey. This was done so that only people who received an invitation link could complete the survey. When respondents clicked on the link and arrived on the RedCap survey page, they saw an institutional review board–approved consent script. After clicking on a button indicating their consent to participate, they were directed to the survey. Participants received a US $20 gift card for completing the survey. Networks were constructed in April 2016. Surveys were sent out from May 2016 to June 2018.

Measures

Participants self-reported their age, sex, race and ethnicity, and education. Social network characteristics were assessed by asking participants how many Twitter users they followed and how many Twitter users followed them. The survey asked which of the following products the participants had used in the past month: cigarettes, e-cigarettes, cigars, pipe tobacco, blunts, hookah, smokeless tobacco, cigarillo, marijuana, and alcohol.

Statistical Analysis

The 3 groups (opinion leaders, followers, and general users) were compared on all measures, using analysis of variance for normally distributed continuous variables, the Kruskal-Wallis test for nonnormally distributed continuous variables, and chi-square for categorical variables.

Results

Demographic Differences Across Groups

The sample included 1433 completed surveys—347 opinion leaders, 567 followers, and 519 general users. The followers (mean age 22.4 years) were significantly younger than the opinion leaders (mean age 24.2 years) and the general users (mean age 25.2 years), $P<.001$. Compared with opinion leaders and general users, followers were more likely to be Hispanic ($P=.03$). General users were more likely than opinion leaders and followers to be African American ($P=.03$). Compared with opinion leaders and general users, followers had less education: only 11% (40/380) of followers had a bachelor's degree or higher as compared with 19% (45/241) of opinion leaders and 26% (92/350) of general users ($P<.001$).

Tobacco or Nicotine Product and Other Substance Use

For most of the tobacco products, opinion leaders reported the highest past month use prevalence, followed by followers and general users (Table 1). This pattern was evident for cigarettes, e-cigarettes, cigars, blunts, hookah, and cigarillos. Opinion leaders had the highest nicotine dependence scores, followed by followers and general users. Opinion leaders were also more likely than followers and general users to have used alcohol or marijuana in the past month.

XSL•FO
RenderX

Table 1. Comparison of opinion leaders, followers, and general users.

Twitter user characteristics	Opinion leaders (n=347)	Followers (n=567)	General users (n=519)	*P* value
Social network size				
Number of Twitter users who follow the respondent (median)	1000	554	503	.001
Number of Twitter users whom the respondent follows (median)	428	375	366	.01
Age (years)	24.2	22.4	25.2	<.001
Female, n/N (%)	122/242 (50.4)	211/392 (53.8)	202/360 (56.1)	.39
Race and ethnicity, n/N (%)				
African American	26/242 (11)	40/392 (10)	67/360 (19)	.03
Asian or Pacific Islander	15/242 (6)	21/392 (5)	14/360 (4)	.03
Hispanic	51/242(21)	102/392 (26)	74/360 (21)	.03
White	129/242 (53)	201/392 (51)	176/360 (49)	.03
Other	21/242 (9)	28/392 (7)	29/360 (8)	.03
Education, n/N (%)				
High school or less	85/241 (35)	155/380 (41)	109/350 (31)	<.001
Some college	111/241 (46)	185/380 (49)	149/350 (43)	<.001
Bachelor's degree or higher	45/241 (19)	40/380 (11)	92/350 (26)	<.001
Number of tobacco products used in past month (mean)	1.31	.98	.66	<.001
Use of specific products in the past month, n/N (%)				
Cigarettes	93/347 (27)	117/567 (21)	83/519 (16)	.01
E-cigarettes	65/347(19)	93/567 (16)	50/519 (10)	.01
Cigar	30/34 (9)	37/567 (7)	23/519 (4)	.04
Pipe	37/347 (11)	64/567 (11)	28/519 (6)	.01
Blunt	96/347 (28)	118/567 (21)	73/519 (14)	.01
Hookah	45/347 (13)	51/567 (9)	29/519 (6)	.01
Smokeless	13/347 (4)	7/567 (1)	11/519 (2)	.04
Cigarillo	60/347 (17)	67/567 (12)	41/519 (8)	.01
Other substance use in the past month, n/N (%)				
Alcohol	159/229 (69)	221/378 (58)	215/354 (61)	.02
Marijuana	103/219 (47)	142/371 (38)	96/330 (29)	<.001
Nicotine dependence score (mean)	1.48	1.18	1.02	.04
Number of tobacco brands followed on Twitter (mean)	.23	.18	.13	.11

Discussion

Principal Findings

Findings suggest that opinion leaders were more likely to report past month use of tobacco products than their followers; followers, in turn, were more likely to report past month use of these products than general Twitter users. The followers appeared to be an especially vulnerable group; they tended to be younger and have lower education. Opinion leaders had higher nicotine dependence scores and were more likely to report past month alcohol or marijuana use compared with followers and general users.

Tobacco opinion leaders on Twitter use a wide variety of tobacco products and other substances. As opinion leaders are typically held in high esteem by their followers, they play an important role in establishing and conveying social norms [19]. Opinion leaders who discuss their polytobacco and polysubstance use on Twitter might lead their followers to believe that these behaviors are normative, safe, or socially admirable. Followers, in turn, might emulate opinion leaders' levels of tobacco use and become nicotine dependent.

Social media–based tobacco campaigns can address tobacco use disparities by tailoring messages that resonate with followers. Such focused efforts can potentially play an important role in educating followers who are typically younger and less

XSL•FO
RenderX

educated than the other groups. Past evidence suggests that network-based interventions that involve identifying peer messengers result in improved health behaviors and more targeted delivery of interventions [13,20]. Social network analysis of a social media–based intervention also revealed that participants from vulnerable demographic groups (younger youth and females) may require additional outreach efforts [21]. Future tobacco health communication campaigns can take advantage of strategic delivery of health messages to followers on social media.

Solely using Twitter metrics to identify opinion leaders can be misleading, as bots, celebrities, and other nontraditional accounts can distort actual user classifications; Twitter metrics such as the number of followers are more likely to measure popularity rather than influence [16,17]. By using social network analysis in combination with Twitter metrics in this study, we are able to systematically identify emergent clusters in the Twitter tobacco network and then apply Twitter metrics to identify subgroup opinion leaders. This method helps prevent over-reliance on Twitter metrics such as follower or retweet count as the sole metric to define opinion leadership.

Limitations

The study utilized a convenience sample by sending unsolicited messages to Twitter users. The tobacco opinion leaders, followers, and general Twitter users were selected on the basis of their positions in the Twitter social network; they had not previously expressed interest in participating in surveys. Twitter users who read their direct messages, click on a survey link, and complete a Web-based survey might not be representative of the general Twitter population; in addition, we have no method to verify that the user who takes the survey is the same as the original Twitter user who received the link.

Conclusions

Despite these limitations, these findings provide important new information about people who disseminate and receive tobacco-related information on Twitter. Opinion leaders are influential as they occupy central positions in the social network and have the potential to communicate with a wide audience of Twitter users. Our findings indicate that tobacco opinion leaders use a wide variety of tobacco products as well as other substances. They may disseminate these attitudes to their Twitter followers who tend to be members of vulnerable populations (eg, young adults and lower educational attainment). Over time, repeated exposure to messages from tobacco opinion leaders could place followers at an increased risk for tobacco product experimentation and escalation. Although this survey was restricted to Twitter users aged 18 years and older, it is likely that younger Twitter users also follow tobacco opinion leaders, and these opinion leaders' messages could persuade them to experiment with tobacco. This study demonstrates that it is possible to identify tobacco opinion leaders on Twitter and their followers and opens the opportunity to apply other methods of supplementing Twitter measures to classify Twitter users. Opinion leaders on the Web have large, well-connected social networks of social media users who may look to them for information, opinions, and advice. If the information disseminated by opinion leaders on the Web is incorrect or biased, their followers could make important decisions based on faulty information. Future research should determine how opinion leaders influence their followers' offline tobacco behaviors.

Acknowledgments

Research reported in this publication was supported by grant #P50CA180905 and U54CA180905 from the National Cancer Institute and the Food and Drug Administration (FDA) Center for Tobacco Products. The National Institutes of Health (NIH) or FDA had no role in the study design, collection, analysis and interpretation of data, writing the report, and decision to submit the report for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or FDA.

Authors' Contributions

All authors contributed to writing the manuscript and approved the final draft. KC developed the algorithm for constructing the Twitter social networks and identifying opinion leaders, followers, and general Twitter posters. AM conducted data analysis and contributed to the interpretation of findings. DS contributed to conceptualization of the study, oversaw data collection, and contributed to the interpretation of the findings. JPA contributed to conceptualization of the study and interpretation of findings. TBC contributed to conceptualization of the study and interpretation of findings. JU contributed to conceptualization of the study, data analysis, and interpretation of findings.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of tobacco-related keywords.

[DOCX File, 32KB - jmir_v21i6e12676_app1.docx]

References

XSL•FO
RenderX

1. Ayers JW, Leas EC, Allem J, Benton A, Dredze M, Althouse BM, et al. Why do people use electronic nicotine delivery systems (electronic cigarettes)? A content analysis of Twitter, 2012-2015. PLoS One 2017;12(3):e0170702 [FREE Full text] [doi: 10.1371/journal.pone.0170702] [Medline: 28248987]
2. Cole-Lewis H, Pugatch J, Sanders A, Varghese A, Posada S, Yun C, et al. Social listening: a content analysis of e-cigarette discussions on Twitter. J Med Internet Res 2015;17(10):e243 [FREE Full text] [doi: 10.2196/jmir.4969] [Medline: 26508089]
3. Zhan Y, Liu R, Li Q, Leischow SJ, Zeng DD. Identifying topics for e-cigarette user-generated contents: a case study from multiple social media platforms. J Med Internet Res 2017 Jan 20;19(1):e24 [FREE Full text] [doi: 10.2196/jmir.5780] [Medline: 28108428]
4. Kirkpatrick MG, Cruz TB, Goldenson NI, Allem J, Chu K, Pentz MA, et al. Electronic cigarette retailers use Pokémon Go to market products. Tob Control 2017 Dec;26(e2):e145-e147. [doi: 10.1136/tobaccocontrol-2016-053369] [Medline: 28044010]
5. Allem J, Cruz TB, Unger JB, Toruno R, Herrera J, Kirkpatrick MG. Return of cartoon to market e-cigarette-related products. Tob Control 2018 Jul 26:-. [doi: 10.1136/tobaccocontrol-2018-054437] [Medline: 30049688]
6. Allem J, Chu K, Cruz TB, Unger JB. Waterpipe promotion and use on Instagram: #hookah. Nicotine Tob Res 2017 Jan 11;19(10):1248-1252. [doi: 10.1093/ntr/ntw329] [Medline: 28077449]
7. Kostygina G, Tran H, Shi Y, Kim Y, Emery S. 'Sweeter Than a Swisher': amount and themes of little cigar and cigarillo content on Twitter. Tob Control 2016 Oct;25(Suppl 1):i75-i82 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053094] [Medline: 27697951]
8. Chu K, Allem J, Cruz TB, Unger JB. Vaping on Instagram: cloud chasing, hand checks and product placement. Tob Control 2016 Sep 22:575-578. [doi: 10.1136/tobaccocontrol-2016-053052] [Medline: 27660111]
9. Unger JB, Escobedo P, Allem J, Soto DW, Chu K, Cruz T. Perceptions of secondhand e-cigarette aerosol among Twitter users. Tob Regul Sci 2016 Apr;2(2):146-152 [FREE Full text] [doi: 10.18001/TRS.2.2.5] [Medline: 28090560]
10. Depue JB, Southwell BG, Betzner AE, Walsh BM. Encoded exposure to tobacco use in social media predicts subsequent smoking behavior. Am J Health Promot 2015;29(4):259-261. [doi: 10.4278/ajhp.130214-ARB-69] [Medline: 24670071]
11. Wu S, Hofman J, Mason W, Watts D. Who says what to whom on Twitter. 2011 Presented at: International World Wide Web Conference; March 28-April 1, 2011; Hyderabad, India p. 705-714 URL: http://www.wwwconference.org/proceedings/www2011/proceedings/p705.pdf
12. Park CS, Kaye BK. The tweet goes on: interconnection of Twitter opinion leadership, network size, and civic engagement. Comput Hum Behav 2017 Apr;69:174-180. [doi: 10.1016/j.chb.2016.12.021]
13. Valente TW. Network interventions. Science 2012 Jul 6;337(6090):49-53. [doi: 10.1126/science.1217330] [Medline: 22767921]
14. Feick L, Price LL. The market maven: a diffuser of information marketplace. J Market 1987 Jan;51(1):83-97. [doi: 10.2307/1251146]
15. Soneji S, Yang J, Knutzen KE, Moran MB, Tan AS, Sargent J, et al. Online tobacco marketing and subsequent tobacco use. Pediatrics 2018 Jan 2;141(2):e20172927. [doi: 10.1542/peds.2017-2927]
16. Riquelme F, González-Cantergiani P. Measuring user influence on Twitter: a survey. Inf Process Manag 2016 Sep;52(5):949-975. [doi: 10.1016/J.IPM.2016.04.003]
17. Garcia D, Mavrodiev P, Casati D, Schweitzer F. Understanding popularity, reputation, and social influence in the Twitter society. Policy Intern 2017;9(3):64. [doi: 10.1002/poi3.151]
18. Newman MEJ. Modularity and community structure in networks. Proc Natl Acad Sci U S A 2006 Jun 6;103(23):8577-8582 [FREE Full text] [doi: 10.1073/pnas.0601602103] [Medline: 16723398]
19. Weeks BE, Ardèvol-Abreu A, Gil de Zúñiga H. Online influence? Social media use, opinion leadership, and political persuasion. Int J Public Opin Res 2015 Dec 31:edv050. [doi: 10.1093/ijpor/edv050]
20. Rice E, Yoshioka-Maxwell A, Petering R, Onasch-Vera L, Craddock J, Tambe M, et al. Piloting the use of artificial intelligence to enhance HIV prevention interventions for youth experiencing homelessness. J Soc Social Work Res 2018 Dec;9(4):551-573. [doi: 10.1086/701439]
21. Rice E, Tulbert E, Cederbaum J, Barman Adhikari A, Milburn NG. Mobilizing homeless youth for HIV prevention: a social network analysis of the acceptability of a face-to-face and online social networking intervention. Health Educ Res 2012 Jan 13;27(2):226-236. [doi: 10.1093/her/cyr113]

Abbreviations

API: application programming interface
FDA: Food and Drug Administration
NIH: National Institutes of Health

XSL•FO
RenderX

Edited by G Eysenbach; submitted 02.11.18; peer-reviewed by G Kostygina, A Benis, R Theis; comments to author 24.01.19; revised version received 20.03.19; accepted 27.04.19; published 04.06.19.

Please cite as:
Chu KH, Majmundar A, Allem JP, Soto DW, Cruz TB, Unger JB
Tobacco Use Behaviors, Attitudes, and Demographic Characteristics of Tobacco Opinion Leaders and Their Followers: Twitter Analysis
J Med Internet Res 2019;21(6):e12676
URL: https://www.jmir.org/2019/6/e12676/
doi:10.2196/12676
PMID:31165716

©Kar-Hai Chu, Anuja Majmundar, Jon-Patrick Allem, Daniel W Soto, Tess Boley Cruz, Jennifer B Unger. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adolescent Perspectives on the Use of Social Media to Support Type 1 Diabetes Management: Focus Group Study

Faisal S Malik[1,2], MD, MSHS; Neil Panlasigui[2], BS; Jesse Gritton[2], MPH; Harsimrat Gill[2], BS; Joyce P Yi-Frazier[3], PhD; Megan A Moreno[4], MD, MPH

[1]Division of Endocrinology and Diabetes, Department of Pediatrics, University of Washington, Seattle, WA, United States

[2]Center for Child Health, Behavior, and Development, Seattle Children's Research Institute, Seattle, WA, United States

[3]Center for Clinical and Translational Research, Seattle Children's Research Institute, Seattle, WA, United States

[4]Division of General Pediatrics and Adolescent Medicine, Department of Pediatrics, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Faisal S Malik, MD, MSHS
Division of Endocrinology and Diabetes
Department of Pediatrics
University of Washington
4800 Sand Point Way NE
M/S OC 7 820, PO Box 5371
Seattle, WA, 98105
United States
Phone: 1 206 987 0121
Fax: 1 206 987 2720
Email: faisal.malik@seattlechildrens.org

Abstract

Background: A majority of adolescents report the use of some form of social media, and many prefer to communicate via social networking sites. Social media may offer new opportunities in diabetes management, particularly in terms of how health care teams provide tailored support and treatment to adolescents with diabetes.

Objective: The aim of this study was to explore the experiences and perspectives of adolescents with type 1 diabetes on the feasibility of social media use as a tool to collaboratively manage their diabetes with their diabetes care team.

Methods: Focus groups of adolescents with type 1 diabetes were conducted in the Seattle metropolitan area in Washington State. Semistructured questions were used to elicit views around the preferred means of communication with the adolescents' diabetes care team, how to best support diabetes self-management, and how social media could be used outside of the clinic setting by the diabetes care team to engage with adolescents with type 1 diabetes. Focus groups were audio recorded and transcribed verbatim. Qualitative content analysis was carried out, and emergent themes were subsequently mapped onto 4 domains of feasibility, which included acceptability, demand, implementation, and practicality.

Results: Participants included 45 adolescents with type 1 diabetes (mean age 15.9, SD 1.7 years; 58% male; diabetes duration mean 6.2, SD 3.6 years; 76% on insulin pumps; 49% wore continuous glucose monitors; 93% reported use of social media; 84% used smartphones as the primary means for social media access). A total of 7 major topics were identified and mapped onto areas consistent with our focus on feasibility. For acceptability and demand, participants expressed how communication over social media could help facilitate (1) improved communication outside of clinic visits to optimize diabetes management, (2) independence in diabetes self-management, (3) connection to other youth with diabetes for additional diabetes support, and (4) delivery of more timely and personalized care. Addressing implementation and practicality, participants shared the need to (1) ensure patient privacy, (2) maintain professional nature of provider-patient relationship, and (3) recognize that social media is not currently used for medical care by youth with diabetes.

Conclusions: Adolescents with type 1 diabetes expressed interest in the use of social media as a tool to support diabetes management and increase engagement with their diabetes care team. Specific implementation measures around privacy and professionalism should be considered when developing a social media intervention to facilitate communication between adolescents and care teams.

XSL•FO
RenderX

(J Med Internet Res 2019;21(6):e12149) doi:10.2196/12149

KEYWORDS

type 1 diabetes; adolescents; social media; qualitative research

Introduction

Background

Despite significant advances in medical treatment, less than 20% of adolescents with type 1 diabetes meet recommended targets for glycemic control [1]. Poor diabetes self-care adherence and suboptimal glycemic control place adolescents with type 1 diabetes at risk for decreased quality of life because of acute (eg, diabetic ketoacidosis) and chronic (eg, blindness, kidney failure) complications [2,3]. Novel strategies to support and improve adolescent engagement and adherence to type 1 diabetes self-care are needed [2,4].

The use of social media may offer one avenue to better meet the needs of adolescents with type 1 diabetes outside the ambulatory care setting. Currently, telephone, email, regular mail, and patient portals are the primary means of communication between health care teams and patients [5,6]. However, these traditional platforms show low rates of acceptability and engagement among adolescents [7]. In contrast, social media is being used by a majority of adolescents, and many cite a preference for communication via social networking sites (SNS). Of those using social media, approximately 90% of teens share that they are on the Web several times a day or almost constantly [8].

Potential for Social Media Use in Diabetes Management

The evolution of social media over the past decade now offers adolescents an array of potential benefits, including access to information, social support, and far-reaching communication tools that have the potential to facilitate diabetes management support outside of the ambulatory care setting. Specifically, the unique functions and features of different social media platforms provide youth a variety of perceived affordances (ie, properties of different social media apps that can be recognized by users and that can contribute to their function) to support chronic disease management [9]. For example, many adolescents with type 1 diabetes may believe that certain features of social media offer cognitive affordances, such as the ability to expand one's learning about type 1 diabetes and how to improve self-care [10]. Other adolescents may be drawn to features of certain social media platforms that offer social affordances, including a sense of belonging to a group related to diabetes, given the interactive nature of social media [10]. These perceived affordances highlight the potential for social media to be used to provide Web-based diabetes education, peer support, and decision advice through real-time bidirectional communication in a digital environment [11]. However, before such affordances can be explored or realized, an understanding of adolescents' experiences and perspectives regarding the feasibility of social media tools is required.

Adolescents with diabetes, in particular, appear to be ideal candidates for chronic disease management support using social media, given their extensive use of technology as part of their daily diabetes management, including devices that monitor glucose and deliver insulin, diabetes-specific apps on smartphones, and internet-enabled education and support programs [12-14]. Using more mainstream SNS that are already well integrated into adolescent daily lives may be a more suitable option for supporting diabetes management outside of the clinic setting, particularly as adolescents routinely use social media to obtain health-related information, as well as social and emotional support [15-17]. Our primary aim was to explore adolescents' experiences and perspectives on the feasibility of using social media to collaboratively manage their type 1 diabetes with their diabetes care team.

Methods

Study Design

This qualitative study used focus groups to explore adolescents' views on how social media use and communication with diabetes care teams could potentially support diabetes management [18]. Focus groups provide the opportunity to acknowledge the participants as the experts. Therefore, the results are likely to inform the development of services that could be more amenable to use by adolescents with type 1 diabetes [19]. In addition, focus groups can be particularly effective with children and adolescents, as the format reduces the pressure for a particular individual to respond, as would be the case in interviews, thereby limiting any adolescent concerns about researcher expectations [19,20]. The format also allows for spontaneous, free-flowing conversations guided by a skilled moderator. Our methods were reported using the Consolidated Criteria for Reporting Qualitative Research guidelines as a framework [21]. The Seattle Children's Research Institute Institutional Review Board approved the study procedures and ensured that ethical principles were applied to research activities.

Study Participants and Recruitment

A purposeful sample of adolescents with type 1 diabetes that received care at Seattle Children's Hospital Diabetes Clinics was recruited from the Seattle metropolitan area in Washington State. English-speaking adolescents (aged 13-19 years) with a type 1 diabetes diagnosis were identified through the medical record and approached via mail and phone recruitment. No current or previous experience with social media was required for participation; our goal was to allow for a full range of viewpoints, including those of adolescents with little or no social media experience.

Participants 18 years and older gave informed written consent before study enrollment. Caregivers provided written consent, and youth participants provided written assent for participants under 18. Before focus group participation, participants completed an anonymous questionnaire to collect demographic data, information about social media use, and diabetes

XSL•FO
RenderX

management. Participants were provided with food at the focus group and compensated US $50 for focus group participation.

Facilitator Guide Development

The research team comprised of a male pediatric endocrinologist with health services and qualitative research training (FM), a female adolescent medicine physician and investigator with extensive social media and qualitative research experience (MM), a female clinical research assistant with social media and qualitative research experience (JG), a male clinical research assistant with a background in patient experience (NP), and a female research health psychologist with social media and diabetes outcomes experience (JYF).

A semistructured guide was developed by the research team to explore optimal communication with an adolescent's diabetes care team on the basis of the key components of the Health Belief Model [22-24]. The diabetes care team was defined as all health care professionals adolescents interact with regarding diabetes management, including diabetes providers, nurses, certified diabetes educators, nutritionists, and social workers.

Focus Groups

Five focus groups with a total of 45 participants were conducted between October and December 2016 in 3 cities in the greater Seattle area (Bellevue, Washington; Everett, Washington; Seattle, Washington). Focus groups were conducted with 8 to 10 participants per group. The Seattle focus groups took place in private conference rooms at the local children's hospital where most youth received their diabetes care; Bellevue and Everett focus groups were held in private meeting rooms at a local library. Focus groups lasted between 110 and 120 min and were audio recorded.

Each focus group was comoderated by FM and JG. NP served as a trained note taker for all focus groups. Open-ended questions and probes were used to encourage a broad discussion about participants' attitudes and opinions around the preferred means of communication with their diabetes care team, how to best support diabetes self-management, and how social media could be used outside of the clinic setting by the diabetes care team to engage with adolescents with type 1 diabetes. The research team conducted debriefs after each focus group, and central themes were documented in focus group summaries. Per standard qualitative methodology regarding saturation, participant recruitment ended when no new themes were identified in the debriefing sessions [25].

Data Analysis

Audio recordings were transcribed verbatim and reviewed by the moderator (FM) for accuracy. Using open coding, members of the research team (FM, NP, and JG) reviewed observer notes and transcripts to identify emergent concepts related to social media communication between adolescents with type 1 diabetes and their diabetes care team to support diabetes management [26]. These concept codes and their definitions were then discussed with the full research team and organized into a codebook.

After assurance of a foundational codebook, transcripts were uploaded to Dedoose version 8.0 software, a Web-based app for managing, analyzing, and presenting qualitative research data [27]. A total of 2 members of the research team (FM, NP) then independently completed line-by-line open coding and applied codes to all transcripts. Questions, disagreements, and newly proposed codes were discussed and resolved through regular bimonthly analysis meetings until consensus was reached. Codes and their corresponding definitions were edited accordingly. Upon completion of all coding, the research team grouped the most significant themes and decided upon the most salient domains, which are presented in this paper. These themes were then mapped onto 4 areas of focus commonly explored by feasibility studies, including acceptability, demand, implementation, and practicality [28].

Results

Demographics

The study included 45 adolescents with a mean age of 15.9 (SD 1.7) years. A total of 34 out of 45 participants (76%) reported using an insulin pump, with approximately half of the participants sharing that they were using a continuous glucose monitoring system to support diabetes management. A total of 42 out of 45 participants (93%) were current users of social media, and 38 out of 45 participants (84%) used a smartphone for social media access (see Table 1).

XSL•FO
RenderX

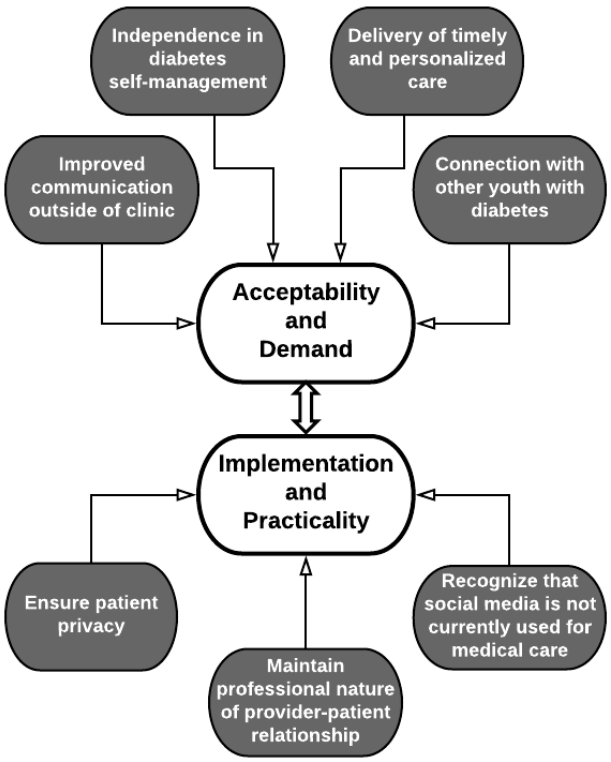
Table 1. Focus group participant demographic data (N=45 adolescents).

Participant characteristics	Values
Age (years), mean (SD)	15.9 (1.7)
Gender (male), n (%)	26 (58)
Race (non-Hispanic white), n (%)	35 (78)
Parental education (college or higher), n (%)	33 (73)
Diabetes duration (years), mean (SD)	6.2 (3.6)
Use insulin pump, n (%)	34 (76)
Use continuous glucose monitor, n (%)	22 (49)
Use social media, n (%)	42 (93)
Social media platform most often used, n (%)	
Snapchat	15 (33)
Facebook	11 (24)
Instagram	8 (18)
Twitter	4 (9)
YouTube	7 (16)
Use smartphone for social media, n (%)	38 (84)

Summary of Findings

A total of 7 major topics were identified consistent with this study focused on feasibility (Figure 1). Adolescent participants expressed that communication over social media could help facilitate (1) improved communication outside of clinic visits to optimize diabetes management, (2) independence in diabetes self-management, (3) connection to other youth with diabetes for additional diabetes support, and (4) delivery of more personalized care. In addition, participants shared important considerations regarding the implementation and practicality of developing a social media intervention for youth with diabetes, including the need to (1) ensure patient privacy, (2) maintain professional nature of provider-patient relationship, and (3) recognize that social media is not currently used for medical care by youth with diabetes.

Figure 1. Key feasibility domains for social media use with diabetes care team.



XSL•FO
RenderX

Domain 1: Acceptability and Demand

Adolescents in this study provided insights into the potential acceptability and demand for social media communication as a tool to support diabetes management and promote engagement with their diabetes care team. Participants also expressed their views regarding the suitability of and comfortability around using social media in this capacity.

Improved Communication Outside of Clinic Visits to Optimize Diabetes Management

For intermittent contact with the care team between clinic visits, focus group participants expressed a preference for communication over social media, compared with existing methods, such as phone or email. In comparison to email, youth stated that social media is more widely used, checked more often, and more user friendly on a smartphone. Participants highlighted the multifaceted potential of social media, noting that in addition to direct private messaging, it offers an opportunity to share diabetes-related educational content, news updates, and current research through visual means.

> *I don't really check my email a ton. Social media would probably be the easiest way to get a hold of us people.* [Adolescent 3, FG1]

> *I would also say that it's very convenient because a lot of social media [platforms] are on your phone…usually everybody has their phone on them, especially as an adult, and it's just at the tip of your fingers. It would be a lot easier to respond…I can't stand emailing people through my phone. I have to do it on a computer, and that's just my personal preference, but I feel like a lot of people are like that.* [Adolescent 5, FG5]

Adolescents from the focus groups also shared that a designated diabetes care team contact that could be reached via social media would likely be easier to access, as they would not have to experience lengthy wait times to be triaged to the appropriate diabetes care team member as they do now when contacting the team over the phone.

> *I feel like I have to jump through so many hoops to get to [talk to a diabetes care team member]. The transfers and the call, or just them not being there, and just not even wanting to call because I'll know I'll have to do that.* [Adolescent 1, FG2]

Participants also shared that social media communication is less formal for them than having to write an email, and thus it would offer a mode of communication that is more conducive for submitting questions and brief check-ins between clinic visits. In addition, they perceived email as being overwhelmingly cluttered and felt that social media offered a more organized platform for communication.

> *I get a lot of emails and it's really hard to distinguish what is spam and what is not [since] sometimes it doesn't go into the spam box. I'll click on something and, oh my god, it will open up something completely different, which is really complicated, especially if I am looking for an important email and I can't find it. I feel like a social media page would be a lot better because depending on how it would be set up, it would be organized.* [Adolescent 7, FG1]

Independence in Diabetes Self-Management

Youth shared that the opportunity to use social media offered a means of empowering them to be more responsible for their diabetes self-care and potentially facilitating increased autonomy.

> *My parents are normally the ones who would talk with the doctors and stuff and so with social media it would be like I'm taking more control over what's happening…it's easier to talk with the doctors and stuff instead of having to call them, find the phone number and all that stuff. You can just send them a message.* [Adolescent 4, FG1]

Many felt that engaging their diabetes care team through social media offered an avenue to reach out to the medical team without the help of their caregivers and better supported continued correspondence over time.

> *Usually when my doctor is communicating with someone in my family about my diabetes, it's with my mother…[But] my mom doesn't know everything that is going on during the school day or hanging out with friends with my diabetes, so it would be nice…[if] I could talk to [my provider] about what's going on in my personal life revolving around diabetes.* [Adolescent 5, FG1]

In addition, participants shared that having a more direct line of communication with the diabetes care team would be beneficial, particularly for adolescents preparing for a transition to college or into the workforce. Several students graduating from high school highlighted that social media, in their opinion, could allow for a smoother transition by increasing their access to convenient remote support, especially as they attempt to navigate their diabetes self-care in a new environment without caregiver support.

Delivery of Timely and Personalized Diabetes Care

Participants highlighted that they were interested in engaging with their diabetes care team on social media, as they felt it would provide an opportunity to follow-up on self-management goals between visits. Although most adolescents shared that they do not communicate between clinic visits, many participants felt that through social media, they would potentially be able to meaningfully communicate about diabetes management goals by touching base at regular short intervals between clinic visits.

> *I know we have the goals we set at our check ups, or at our appointments, but… figuring out a short term goal for yourself, and then being like okay, we're gonna check back two weeks from now [over social media].* [Adolescent 4, FG2]

In addition, adolescents expressed that the option of communicating over social media could also allow them to address acute issues that may arise between visits.

> *If it's a question that affects your care during those three months [between clinic visits], or it does*

XSL•FO
RenderX

> *something to help what you're doing within diabetes, then that could be helpful because you get the answer then and not have to wait.* [Adolescent 2, FG2]

Several youth felt that their relationship with the care team could potentially be enhanced by knowing who the care team members were outside of the clinic setting, which they felt could lead to increased trust and translate into more open communication. Some were eager to establish closer relationships with their providers outside of the clinic setting, whereas others could at least appreciate the efforts to build better rapport overall.

> *I think you would get a more personal relationship with your doctor [through social media] without it being creepy…you would probably get to know each other a bit more and you would be more knowledgeable about things going on.* [Adolescent 1, FG1]

A few participants felt that a better understanding of their personal life could facilitate more personalized diabetes management recommendations to support self-care. As an example, youth mentioned that a diabetes care team member who could see the kinds of food they ate on their social media account might be able to provide better recommendations regarding insulin dose adjustments and nutrition recommendations.

> *[Through social media] they would know more about me, and what I like and what kind of foods I like, so that could also factor into insulin pump settings.* [Adolescent 8, FG4]

Connection to Other Youth With Diabetes to Help Optimize Diabetes Management

Most participants shared that they do not currently have a circle of friends with whom they can relate to regarding their chronic disease, and although some may know of a peer or family member who is also living with diabetes, relationships built around diabetes are not common. A majority of adolescents agreed that emotional support from peers would be valuable. Thus, many focus group participants expressed an interest in the idea of being able to engage over social media with not only their diabetes care team but also other youth with diabetes.

> *You can form really strong bonds over the internet and I think if you have something as big as diabetes in common then like you could probably bond really fast. I mean I trust you guys and I've only met you today.* [Adolescent 4, FG1]

Some participants viewed social media as an avenue for connecting with different resources and gaining new perspectives on diabetes care. These participants highlighted the potential of exchanging ideas with their peers on the Web and gaining insight into their unique care routines. Participants felt that unlike email, which limits their ability to conveniently gain input from other youth, social media would facilitate easier engagement with a larger network of peers. Several participants mentioned that simply having a credible and trustworthy social media page related to diabetes management that adolescents could access would provide the means to receive useful diabetes-related information with the freedom to engage as needed.

> *There should be a part on the [social media] page where...so you are seeing different perspectives, like from someone who lives on the east coast and west coast or someone who lives where there is a lot of resources and not a lot of resources. To see what is working and what's not working because for me that would kind of open my eyes to things that are out there, what works and what doesn't work and for me I like trying new things so it would be pretty cool to see…their daily routine.* [Adolescent 2, FG1]

> *If you were talking to someone that had a certain way they handled their diabetes, that you liked...you could exchange tips and tricks for treating diabetes.* [Adolescent 3, FG5]

Furthermore, adolescent participants felt that the ability to use social media to support diabetes management presented a unique opportunity to build a peer community without parental involvement, and it could help them in embracing how they are different from other youth who do not have diabetes.

> *I feel like having a community is a huge step towards really accepting it because I feel I haven't really accepted it yet, even though it's been like a really long time. It's really hard to get used to, and there's always ups and downs and different variables it affects. It just changes all the time, and there's no norm, so it's really hard to get the feel of that. I feel, like, having other people who can be a different norm with you would be a big step in helping you improve your attitude towards diabetes, and make it fun.* [Adolescent 5, FG5]

Domain 2: Implementation and Practicality

Participants also shared their views on elements related to successful implementation for using social media to communicate with their diabetes care team. Specifically, adolescents highlighted items that should be considered during intervention development to support increased comfort with social media use for diabetes management.

Ensure Patient Privacy

When discussing measures that would need to be taken to promote buy-in from youth with diabetes to use social media to support diabetes management, youth cited the need to safeguard private social media conversations. For example, youth highlighted they did not want their direct messages (DM) with the diabetes care team from becoming public.

> *In a DM [conversation] it's only going to me ... [but] I feel like if it was me and [care team member] in a DM, I could add anyone at any time, or [the care team member] could add anyone at any time, and then it's no longer private.* [Adolescent 3, FG2]

A majority of participants were not worried that social media may not be secure enough to guarantee the confidentiality of sensitive information. Participants who did express concern were primarily worried that their condition would accidentally

XSL•FO
RenderX

be broadcast on social media, especially to people they had not intended to share diabetes-related information publicly. Youth who did not want to participate in social media for this reason preferred the idea of texting with the diabetes care team rather than using existing methods, such as email or phone.

Maintain Professional Nature of Provider-Patient Relationship

Most youth specifically valued the professional nature of their relationship with their provider, which some felt could be compromised through the use of social media. Several did not want diabetes care team members commenting publicly on pictures, hemoglobin A_{1c} values, or aspects of their personal lives that may be affecting their diabetes management.

> *I think it would be good in some ways, but it might be awkward at the same time, like, if your doctor's following you on Instagram and they can see everything you're posting and all that stuff.* [Adolescent 4, FG3]

Recognize That Social Media is Not Currently Used for Medical Care by Youth With Diabetes

Participants reported that most use social media to share information about their personal lives or about things they find interesting that are usually not related to their diabetes. Although a majority of participants were avid social media users, they shared that the lack of experience using social media to supplement their diabetes management might initially be a barrier to engagement. In addition, they shared they would have to become comfortable with using social media for purposes other than social networking.

> *I think it would be a little weird [engaging over social media] at first. I don't really know, I'm not very creative in thinking of how that would work but, I mean, I would do what they thought of.* [Adolescent 6, FG1]

Discussion

Principal Findings

This qualitative study found that adolescents with type 1 diabetes are interested in using social media as a tool to support diabetes management and increase engagement with their diabetes care team. However, participants also expressed some concern about certain features that are part and parcel of social media platforms. In addition to highlighting social media's strengths as a means of communication for youth, adolescents also shared that having the option to communicate over social media has the potential to promote autonomy and facilitate timely and individualized care. Recognizing that adolescents with type 1 diabetes are at high risk for poor health outcomes [29], social media may offer a means to enhance collaboration with the care team to improve diabetes management outside of the clinic setting.

Collaborative communication between providers and patients is viewed as a key element in achieving favorable health outcomes in disease management [30]. However, challenges around adequate communication and the ability to access timely health-related information continue to be a barrier to optimal type 1 diabetes care for youth [31]. The use of technology has been recommended as an avenue for improving communication with adolescents [30]. Youth in this study provided insights into their specific communication preferences on the basis of the perceived affordances of social media, and they expressed a clear interest in using some features of social media platforms to facilitate increased contact outside of clinic visits. Adolescents with type 1 diabetes viewed social media as an avenue for more direct and timely communication with their diabetes care team than existing means. These results are consistent with trends that highlight increased use of social media communication in adolescents compared with email and voice calls [32].

In an era of increasing electronic health implementation, health care organizations are making use of patient portals to enhance patient-provider communication and support self-management [33]. Specifically, these electronic systems are being designed to allow patients to view their electronic health record and offer options to communicate with providers through secure messaging, refill medications, and schedule appointments [34,35]. However, patients report many usability barriers, such as unfamiliarity with portal features and unknown or lost log-in information [36,37]. Several studies also reveal racial and ethnic disparities in portal enrollment [33,38,39]. In this study, participants reported a high level of familiarity and comfortability with the current social media platforms and perceived many of them to be well designed for smartphone use. Given this, social media may be a more ideal method for adolescents to communicate with their diabetes care team, as it bypasses new design- and navigation-related burdens and is already well accepted and highly used among adolescents [40].

The shared responsibility model is encouraged in diabetes management, and it highlights that maturing adolescents should work toward an increasing level of responsibility in diabetes management while maintaining appropriate support from their caregiver [41]. However, there are few distinct pathways to facilitate the increase in independence for adolescents. This study's results indicate that the availability of social media communication could be a novel means of supporting the adolescent in assuming responsibility for between-visit communication. Without the help of their caregivers, a majority of adolescents stated they would be willing to initiate direct communication with the diabetes care team about nonemergent issues related to their diabetes through social media, something they do not routinely do through the existing communication means available to them. Although caregiver involvement is considered an integral part of supporting adolescent diabetes management, the development of adolescent autonomy has been identified as an important component for successful diabetes management in adulthood, as well as overall psychosocial development in adolescents with type 1 diabetes [42,43]. Providing adolescents with an opportunity to be more independent through social media communication may also improve engagement in self-care, as higher self-reported autonomy in diabetes management has been shown to be correlated with better adherence to treatment in adolescents with type 1 diabetes [42].

XSL•FO
RenderX

Adolescents in this study felt that communicating with their diabetes care team via social media could also enhance personalized care by allowing diabetes care team members to gain a better sense of their lifestyle, engage in more short-term goal setting, and build better rapport between them. The suggestions offered by the participants reflect a desire for patient-centered communication and support, encompassing aspects such as thorough information exchange, collaborative goal setting, and shared decision making [44,45]. These strategies have been shown to be associated with better adherence to treatment and perceptions of empowerment in youth with type 1 diabetes, as they can facilitate a better understanding of the unique needs and preferences of adolescents and encourage active patient engagement [44,46].

The role of social media has expanded to health education and management in youth, with many seeking information on disease management, as well as social and emotional support on the Web [15,17]. This study's results confirm the desire of adolescents with type 1 diabetes to obtain information related to health management through social media, not just from their diabetes care team but also from other youth with type 1 diabetes. However, as recognized by health care professionals themselves, the overall reliability and accuracy of health information accessible to patients on social media can be poor [47,48]. Involving health care professionals in a moderating role in Web-based communication among peers can assist in countering misinformation through distribution of evidence-based medical expertise and facilitate social support for patients [49]. In addition, moderators can encourage patient engagement, maintain focus in discussions, and redirect patients to additional, reliable resources [50].

A barrier to using social media for health communication is the risk for violation of patient privacy. Breaching of patient privacy can carry negative legal implications for the diabetes care team members under violations of Health Insurance Portability and Accountability Act and other professional guidelines [48,51]. In addition, some people are opposed to being portrayed as patients on social media, and they fear the publicizing of their personal health information [52]. Women, in particular, are known to highly value privacy and use restrictive privacy settings at higher rates than men [53]. Participants in this study echoed these attitudes, as they did not desire to share information about diabetes with their larger social media networks. However, participants did acknowledge the use of more private social media features, such as DM, as acceptable means of communication with their care team. Future social media health communication interventions should prioritize patient privacy and leverage social media privacy settings and features to ensure a secure environment.

Most participants in this study expressed a preference for maintaining a professional relationship with their provider. Preservation of professional boundaries on social media has also been recognized as an important issue by the medical community, as health professionals are increasingly using social media to provide evidence-based information to broad communities, network with colleagues, and engage in advocacy or public health initiatives [54,55]. In the efforts to address these concerns, professional medical organizations have published policy statements that include guidelines regarding ethics of professionalism on the Web [56].

Limitations

This study has a few limitations. First, this qualitative study is limited by use of a sample recruited from a major metropolitan area. It is possible that the views of adolescents living in rural settings may differ, even though the option for increased remote care support may be desirable to adolescents with type 1 diabetes, as these patients often face decreased access to clinic facilities and increased transportation challenges. Second, by excluding adolescents who did not speak English, we may have missed themes of particular importance to non-English-speaking adolescents. Finally, given that the majority of the participants were already avid users of social media, this study's results may have shown bias toward an acceptance of social media for health communication purposes. Future studies should aim to explore perceptions regarding social media communication in adolescents of varying racial, ethnic, and socioeconomic backgrounds.

Conclusions

This study's results suggest that the use of certain social media platform features to support diabetes management outside of the ambulatory setting is acceptable to adolescents with type 1 diabetes. Given social media's potential to enhance communication with adolescents, the use of social media as a health collaboration tool among adolescents with type 1 diabetes should be actively considered by diabetes care teams. Although social media offers the potential to improve patient care in a multifaceted manner, elements related to successful implementation should be carefully reviewed, specifically the efforts to preserve privacy and professionalism. In collaboration with technology experts, an understanding of the advantages and disadvantages of the various social media platform features, as well as adolescents' acceptability and enthusiasm for various affordances different platforms of social media offer, should be acquired. Finally, although it is known that a considerable amount of health care professionals already use social media in some fashion related to their field, their perspectives and concerns should be taken into consideration when further exploring the feasibility of social media as a platform for health collaboration with patients.

Acknowledgments

The authors wish to thank Katherine Velasco for her assistance with this study. This work was funded by the Center for Child Health, Behavior and Development at the Seattle Children's Research Institute. Dr Malik's time was supported in part by a K23 Career Development Award from the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (DK119465).

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Foster NC, Beck RW, Miller KM, Clements MA, Rickels MR, DiMeglio LA, et al. State of Type 1 Diabetes Management and Outcomes from the T1D Exchange in 2016-2018. Diabetes Technol Ther 2019 Feb;21(2):66-72. [doi: 10.1089/dia.2018.0384] [Medline: 30657336]
2. Datye K, Moore D, Russell W, Jaser S. A review of adolescent adherence in type 1 diabetes and the untapped potential of diabetes providers to improve outcomes. Curr Diab Rep 2015 Aug;15(8):51 [FREE Full text] [doi: 10.1007/s11892-015-0621-6] [Medline: 26084580]
3. Bryden K, Peveler R, Stein A, Neil A, Mayou R, Dunger D. Clinical and psychological course of diabetes from adolescence to young adulthood: a longitudinal cohort study. Diabetes Care 2001 Sep;24(9):1536-1540. [doi: 10.2337/diacare.24.9.1536] [Medline: 11522695]
4. Rausch J, Hood K, Delamater A, Shroff Pendley J, Rohan J, Reeves G, et al. Changes in treatment adherence and glycemic control during the transition to adolescence in type 1 diabetes. Diabetes Care 2012 Jun;35(6):1219-1224 [FREE Full text] [doi: 10.2337/dc11-2163] [Medline: 22474040]
5. Hausmann J, Touloumtzis C, White M, Colbert J, Gooding H. Adolescent and young adult use of social media for health and its implications. J Adolesc Health 2017 Jun;60(6):714-719 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.12.025] [Medline: 28259620]
6. Coiera E. Communication systems in healthcare. Clin Biochem Rev 2006 May;27(2):89-98 [FREE Full text] [Medline: 17077879]
7. Boogerd E, Kremer J, Prins J, Noordam K, Verhaak C. Adolescents' perspectives on having type I diabetes, on current outpatient diabetes care, and on improvement of care by using the internet. J Participat Med 2016 Aug 24;8:e11 [FREE Full text]
8. Anderson M, Jiang J. Pew Research Center. 2018. Teens, Social Media & Technology 2018 URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2018/05/31102617/PI_2018.05.31_TeensTech_FINAL.pdf [accessed 2018-08-31] [WebCite Cache ID 725hlCYb4]
9. Burgess J, Marwick A, Poell T. The Sage Handbook of Social Media. London: Sage Publications Ltd; 2018.
10. Moreno M, Uhls Y. Applying an affordances approach and a developmental lens to approach adolescent social media use. Digit Health 2019;5:2055207619826678 [FREE Full text] [doi: 10.1177/2055207619826678] [Medline: 30746154]
11. Gabarron E, Årsand E, Wynn R. Social media use in interventions for diabetes: rapid evidence-based review. J Med Internet Res 2018 Dec 10;20(8):e10303 [FREE Full text] [doi: 10.2196/10303] [Medline: 30097421]
12. Giani E, Scaramuzza A, Zuccotti G. Impact of new technologies on diabetes care. World J Diabetes 2015 Jul 25;6(8):999-1004 [FREE Full text] [doi: 10.4239/wjd.v6.i8.999] [Medline: 26240696]
13. Tran J, Tran R, White J. Smartphone-based glucose monitors and applications in the management of diabetes: an overview of 10 salient "apps" and a novel smartphone-connected blood glucose monitor. Clinical Diabetes 2012 Oct 15;30(4):173-178. [doi: 10.2337/diaclin.30.4.173]
14. Frøisland DE, Arsand E, Skårderud F. Improving diabetes care for young people with type 1 diabetes through visual learning on mobile phones: mixed-methods study. J Med Internet Res 2012 Aug 6;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
15. Nordfeldt S, Angarne-Lindberg T, Nordwall M, Ekberg J, Berterö C. As facts and chats go online, what is important for adolescents with type 1 diabetes? PLoS One 2013;8(6):e67659 [FREE Full text] [doi: 10.1371/journal.pone.0067659] [Medline: 23805322]
16. Ravert R, Hancock M, Ingersoll G. Online forum messages posted by adolescents with type 1 diabetes. Diabetes Educ 2004;30(5):827-834. [doi: 10.1177/014572170403000518] [Medline: 15510534]
17. Fergie G, Hunt K, Hilton S. Social media as a space for support: young adults' perspectives on producing and consuming user-generated content about diabetes and mental health. Soc Sci Med 2016 Dec;170:46-54 [FREE Full text] [doi: 10.1016/j.socscimed.2016.10.006] [Medline: 27750067]
18. Kitzinger J. Qualitative research. Introducing focus groups. Br Med J 1995 Jul 29;311(7000):299-302 [FREE Full text] [doi: 10.1136/bmj.311.7000.299] [Medline: 7633241]
19. Heary C, Hennessy E. The use of focus group interviews in pediatric health care research. J Pediatr Psychol 2002;27(1):47-57. [doi: 10.1093/jpepsy/27.1.47] [Medline: 11726679]
20. Peterson-Sweeney K. The use of focus groups in pediatric and adolescent research. J Pediatr Health Care 2005;19(2):104-110. [doi: 10.1016/j.pedhc.2004.08.006] [Medline: 15750555]
21. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]

XSL•FO
RenderX

22. Brownlee-Duffeck M, Peterson L, Simonds J, Goldstein D, Kilo C, Hoette S. The role of health beliefs in the regimen adherence and metabolic control of adolescents and adults with diabetes mellitus. J Consult Clin Psychol 1987 Apr;55(2):139-144. [Medline: 3571665]
23. Bond G, Aiken L, Somerville S. The health belief model and adolescents with insulin-dependent diabetes mellitus. Health Psychol 1992;11(3):190-198. [Medline: 1618173]
24. Patino A, Sanchez J, Eidson M, Delamater A. Health beliefs and regimen adherence in minority adolescents with type 1 diabetes. J Pediatr Psychol 2005 Sep;30(6):503-512. [doi: 10.1093/jpepsy/jsi075] [Medline: 16055488]
25. Bowen G. Naturalistic inquiry and the saturation concept: a research note. Qual Res 2008 Feb;8(1):137-152. [doi: 10.1177/1468794107085301]
26. Ulin PR, Tolley E. Qualitative Methods in Public Health: A Field Guide for Applied Research. San Francisco: Jossey-Bass; 2005.
27. Bernhard G, Mahler C, Seidling H, Stützle M, Ose D, Baudendistel I, et al. Developing a shared patient-centered, web-based medication platform for type 2 diabetes patients and their health care providers: qualitative study on user requirements. J Med Internet Res 2018 Dec 27;20(3):e105 [FREE Full text] [doi: 10.2196/jmir.8666] [Medline: 29588269]
28. Bowen D, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
29. Petitti D, Klingensmith G, Bell R, Andrews J, Dabelea D, Imperatore G, SEARCH for Diabetes in Youth Study Group. Glycemic control in youth with diabetes: the SEARCH for diabetes in Youth Study. J Pediatr 2009 Nov;155(5):668-72.e1 [FREE Full text] [doi: 10.1016/j.jpeds.2009.05.025] [Medline: 19643434]
30. Patel N, Datye K, Jaser S. Importance of patient-provider communication to adherence in adolescents with type 1 diabetes. Healthcare (Basel) 2018 Mar 28;6(2):E30 [FREE Full text] [doi: 10.3390/healthcare6020030] [Medline: 29597246]
31. Valenzuela J, Seid M, Waitzfelder B, Anderson A, Beavers D, Dabelea D, SEARCH for Diabetes in Youth Study Group. Prevalence of and disparities in barriers to care experienced by youth with type 1 diabetes. J Pediatr 2014;164(6):e1-75. [Medline: 24582008]
32. Lenhart A. Pew Research Center. 2012. Teens, Smartphones & Texting URL: http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_Teens_Smartphones_and_Texting.pdf [accessed 2018-08-31] [WebCite Cache ID 725KwxXhK]
33. Irizarry T, DeVito Dabbs A, Curran C. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
34. Wade-Vuturo A, Mayberry L, Osborn C. Secure messaging and diabetes management: experiences and perspectives of patient portal users. J Am Med Inform Assoc 2013 May 1;20(3):519-525 [FREE Full text] [doi: 10.1136/amiajnl-2012-001253] [Medline: 23242764]
35. Weingart S, Rind D, Tofias Z, Sands D. Who uses the patient internet portal? The PatientSite experience. J Am Med Inform Assoc 2006;13(1):91-95 [FREE Full text] [doi: 10.1197/jamia.M1833] [Medline: 16221943]
36. Amante D, Hogan T, Pagoto S, English T. A systematic review of electronic portal usage among patients with diabetes. Diabetes Technol Ther 2014 Nov;16(11):784-793. [doi: 10.1089/dia.2014.0078] [Medline: 24999599]
37. Goel M, Brown T, Williams A, Cooper A, Hasnain-Wynia R, Baker D. Patient reported barriers to enrolling in a patient portal. J Am Med Inform Assoc 2011 Dec;18(Suppl 1):i8-12 [FREE Full text] [doi: 10.1136/amiajnl-2011-000473] [Medline: 22071530]
38. Goel M, Brown T, Williams A, Hasnain-Wynia R, Thompson J, Baker D. Disparities in enrollment and use of an electronic patient portal. J Gen Intern Med 2011 Oct;26(10):1112-1116 [FREE Full text] [doi: 10.1007/s11606-011-1728-3] [Medline: 21538166]
39. Sarkar U, Karter A, Liu J, Adler N, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011 May 1;18(3):318-321 [FREE Full text] [doi: 10.1136/jamia.2010.006015] [Medline: 21262921]
40. Lenhart A. Pew Research Center. Teens, Social Media & Technology Overview 2015 URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/04/PI_TeensandTech_Update2015_0409151.pdf [accessed 2018-08-31] [WebCite Cache ID 725KTyKMm]
41. Peters A, Laffel L, American Diabetes Association Transitions Working Group. Diabetes care for emerging adults: recommendations for transition from pediatric to adult diabetes care systems: a position statement of the American Diabetes Association, with representation by the American College of Osteopathic Family Physicians, the American Academy of Pediatrics, the American Association of Clinical Endocrinologists, the American Osteopathic Association, the Centers for Disease Control and Prevention, Children with Diabetes, The Endocrine Society, the International Society for Pediatric and Adolescent Diabetes, Juvenile Diabetes Research Foundation International, the National Diabetes Education Program, and the Pediatric Endocrine Society (formerly Lawson Wilkins Pediatric Endocrine Society). Diabetes Care 2011 Nov;34(11):2477-2485 [FREE Full text] [doi: 10.2337/dc11-1723] [Medline: 22025785]
42. Hanna K, Guthrie D. Adolescents' behavioral autonomy related to diabetes management and adolescent activities/rules. Diabetes Educ 2003;29(2):283-291. [Medline: 12728755]

XSL•FO
RenderX

43. Young M, Lord J, Patel N, Gruhn M, Jaser S. Good cop, bad cop: quality of parental involvement in type 1 diabetes management in youth. Curr Diab Rep 2014;14(11):546 [FREE Full text] [doi: 10.1007/s11892-014-0546-5] [Medline: 25212099]
44. Powell P, Corathers S, Raymond J, Streisand R. New approaches to providing individualized diabetes care in the 21st century. Curr Diabetes Rev 2015;11(4):222-230 [FREE Full text] [Medline: 25901504]
45. Epstein R, Street J. The values and value of patient-centered care. Ann Fam Med 2011;9(2):100-103 [FREE Full text] [doi: 10.1370/afm.1239] [Medline: 21403134]
46. Reynolds A. Patient-centered care. Radiol Technol 2009;81(2):133-147. [Medline: 19901351]
47. De Martino I, D'Apolito R, McLawhorn A, Fehring K, Sculco P, Gasparini G. Social media for patients: benefits and drawbacks. Curr Rev Musculoskelet Med 2017 Mar;10(1):141-145 [FREE Full text] [doi: 10.1007/s12178-017-9394-7] [Medline: 28110391]
48. Surani Z, Hirani R, Elias A, Quisenberry L, Varon J, Surani S, et al. Social media usage among health care providers. BMC Res Notes 2017 Nov 29;10(1):654 [FREE Full text] [doi: 10.1186/s13104-017-2993-y] [Medline: 29187244]
49. Huh J, Marmor R, Jiang X. Lessons learned for online health community moderator roles: a mixed-methods study of moderators resigning from WebMD communities. J Med Internet Res 2016 Dec 8;18(9):e247 [FREE Full text] [doi: 10.2196/jmir.6331] [Medline: 27608721]
50. Huh J, McDonald D, Hartzler A, Pratt W. Patient moderator interaction in online health communities. AMIA Annu Symp Proc 2013;2013:627-636 [FREE Full text] [Medline: 24551364]
51. Ventola C. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
52. Antheunis M, Tates K, Nieboer T. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ Couns 2013 Sep;92(3):426-431. [doi: 10.1016/j.pec.2013.06.020] [Medline: 23899831]
53. Madden M. Pew Internet. 2012. Privacy Management on Social Media Sites URL: http://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2012/PIP_Privacy_management_on_social_media_sites_022412.pdf [accessed 2018-12-03] [WebCite Cache ID 74ORBCIKB]
54. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
55. Gholami-Kordkheili F, Wild V, Strech D. The impact of social media on medical professionalism: a systematic qualitative review of challenges and opportunities. J Med Internet Res 2013 Aug 28;15(8):e184 [FREE Full text] [doi: 10.2196/jmir.2708] [Medline: 23985172]
56. Farnan JS, Sulmasy LS, Worster BK, Chaudhry HJ, Rhyne JA, Arora VM, American College of Physicians Ethics, Professionalism and Human Rights Committee, American College of Physicians Council of Associates, Federation of State Medical Boards Special Committee on Ethics and Professionalism*. Online medical professionalism: patient and public relationships: policy statement from the American College of Physicians and the Federation of State Medical Boards. Ann Intern Med 2013 Apr 16;158(8):620-627. [doi: 10.7326/0003-4819-158-8-201304160-00100] [Medline: 23579867]

Abbreviations

DM: direct messaging
SNS: social networking sites

Edited by G Eysenbach; submitted 08.09.18; peer-reviewed by J Brož, S Oser, M Seid; comments to author 28.10.18; revised version received 23.12.18; accepted 29.03.19; published 30.05.19.

Please cite as:
Malik FS, Panlasigui N, Gritton J, Gill H, Yi-Frazier JP, Moreno MA
Adolescent Perspectives on the Use of Social Media to Support Type 1 Diabetes Management: Focus Group Study
J Med Internet Res 2019;21(6):e12149
URL: https://www.jmir.org/2019/6/e12149/
doi:10.2196/12149
PMID:31199310

©Faisal S Malik, Neil Panlasigui, Jesse Gritton, Harsimrat Gill, Joyce P Yi-Frazier, Megan A Moreno. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 30.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet

XSL•FO
RenderX

Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations of Social Media Use With Physical Activity and Sleep Adequacy Among Adolescents: Cross-Sectional Survey

Sandhya V Shimoga[1*], PhD; Erlyana Erlyana[1*], MD, PhD; Vida Rebello[1*], MSHCA

Department of Health Care Administration, California State University Long Beach, Long Beach, CA, United States

[*]all authors contributed equally

Corresponding Author:
Erlyana Erlyana, MD, PhD
Department of Health Care Administration
California State University Long Beach
1250 Bellflower Blvd
Long Beach, CA, 90840
United States
Phone: 1 562 985 5800
Fax: 1 562 985 5886
Email: erlyana.erlyana@csulb.edu

Abstract

Background: Adolescents' use of social media, which has increased considerably in the past decade, has both positive and negative influences on adolescents' health and health behaviors. As social media is the most prominent communication tool of choice for adolescents, it is important to understand the relationship between the frequency of social media use and health behaviors among this population.

Objective: The objective of our study was to examine the associations between the frequency of social media use and physical activity and sleep adequacy among middle and high school students.

Methods: We used data from the Monitoring the Future survey (2014 and 2015), a nationally representative, annual, cross-sectional survey of American 8th-, 10th-, and 12th-grade students (N=43,994). Health behaviors examined were frequency of vigorous physical activity and frequency of getting 7 hours of sleep (never/seldom, sometimes, and every day/nearly every day). We measured frequency of social media use using a Likert-like scale (never, a few times a year, 1-2 times a month, once a week, or every day). Multivariable generalized ordered logistic regressions examined the association of social media use with different levels of physical activity and sleep. We estimated marginal effects (MEs) for the main independent variable (social media use frequency) by holding all other variables at their observed values.

Results: The study population comprised 51.13% (21,276/42,067) female students, 37.48% (17,160/43,994) from the South, and 80.07% (34,953/43,994) from a metropolitan area, with 76.90% (33,831/43,994) reporting using social media every day. Among physically active students, frequent social media use was associated with a higher likelihood of vigorous daily exercise (ME 50.1%, 95% CI 49.2%-51.0%). Among sedentary students, frequent social media use was associated with a lower likelihood of vigorous daily exercise (ME 15.8%, 95% CI 15.1%-16.4%). Moderately active students who used social media once or twice a month had the highest likelihood of reporting vigorous daily exercise (ME 42.0%, 95% CI 37.6%-46.3%). Among those who normally got adequate sleep, daily social media users were least likely to report adequate sleep (ME 41.3%, 95% CI 40.4%-42.1%). Among those who were usually sleep deprived, daily social media users were more likely to report adequate sleep (ME 18.3%, 95% CI 17.6%-19.0%).

Conclusions: Regular social media use every day was associated with a reinforcement of health behaviors at both extremes of health behaviors, whereas a medium intensity of social media use was associated with the highest levels of physical activity and lowest sleep adequacy among those with moderate health behaviors. Hence, finding an optimal level of social media use that is beneficial to a variety of health behaviors would be most beneficial to adolescents who are in the middle of the health behavior spectrum.

(J Med Internet Res 2019;21(6):e14290) doi:10.2196/14290

KEYWORDS

adolescent; social media; exercise; sleep

XSL•FO
RenderX

Introduction

Background

Over the past decade, regular use of social media by young adults has increased considerably from 89% in 2014 to 97% in 2016 [1]. Villanti and colleagues also reported that young adults used an average of 7.6 social media sites regularly, with 85% of them using 6 or more sites regularly [1]. Adolescents' time spent on social media has also more than doubled from 4.4 hours weekly in 2007 to 11.1 hours in 2011 [2].

Adolescents are more likely than any other age group to use social media. In 2013, approximately 45% of adolescents in the United States used social media sites daily and, among those, 73% used it to connect with their peers [3,4]. This extensive use of social media during the formative years calls into question the extent of its influence on all aspects of development, including physical and mental health. An emerging body of evidence shows both positive and negative influences of social media use on adolescents' health and health behaviors [5-8]. The benefits of social media use include exposure to new ideas and information and raising awareness of current events and issues. The interactive nature of social media can provide opportunities to engage with peers on issues, access support networks, and improve social inclusion; it may also foster healthy eating habits [8]. However, there are notable negative health outcomes associated with social media use among adolescents. Higher levels of social media use among this population are associated with lower levels of participation in sports activity, less happiness, and more socioemotional difficulties [9]. Smartphone or mobile phone use at night is associated with reduced and disrupted sleep patterns among adolescents [6,10-12]. High-frequency social media use at night was significantly associated with perceived insufficient sleep [13]. While a high risk of poor-quality sleep or sleep disturbance is associated with frequent social media use among adolescents [11,12,14-16], there is some indication that sleeplessness precedes excessive media use [17,18]. Youths who connect with peers face-to-face have positive emotional outcomes compared with those who predominantly use social media to connect with peers [19,20]. Mental health conditions including anxiety and depression are exacerbated by excessive social media use among adolescents [15,21,22]. The predominant conclusion reached by studies on the relationship between social media use and health behaviors is that the time spent on social media use supplants time spent otherwise on physical activity or sleeping.

While many studies have focused on excessive use of social media or the internet, evidence is emerging that the relationship between intensity of internet use and health outcomes is not necessarily linear. Some recent studies suggested a U-shaped relationship between internet use frequency and depressive symptoms, where both low and high levels of internet use were associated with higher risks for depression and lower levels of mental well-being [21,23]. These studies examined cumulative use of all types of digital platforms. We did not find studies that specifically explored the relationship between frequency of social media use and health behaviors among adolescents.

Objective

As social media is the most prominent communication tool of choice for adolescents, it is important to understand the relationship between the frequency of social media use and health behaviors among this population. Therefore, in this study, we examined the association between the frequency of social media use and different levels of physical activity and sleep adequacy among middle and high school students in the United States.

Methods

Data Source

Data for the study were from the Monitoring the Future (MTF) survey, which is a nationally representative, cross-sectional, annual survey of 8th-, 10th-, and 12th-grade students in the United States. MTF collects data through self-administered questionnaires in approximately 420 high schools and middle schools in the contiguous United States [24,25]. To obtain a nationally representative sample of schools and students, the survey employs a multistage random sampling design with (1) geographic areas or primary sampling units, (2) schools (or linked groups of schools) within primary sampling units, and (3) students within sampled schools. Schools are asked to participate for 2 years, and schools refusing participation are replaced with similar schools in terms of geographic location, size, and type of school. Student response rates for the survey were over 80% for all grade levels during all the years considered for our study and almost all nonresponse was due to absenteeism.

Data collection procedures were approved by the Institutional Review Board of the University of Michigan and the survey was funded by the National Institute on Drug Abuse. Public-use data files are available from Inter-university Consortium for Political and Social Research. Descriptions of the survey sampling frame, methodology, and response rates are detailed elsewhere [26].

We derived the study sample from MTF surveys conducted in 2014 and 2015. The data consisted of a pooled sample of 43,994 students in the 8th, 10th, and 12th grades who answered the questions on social media use and at least one of the questions on health behaviors. Eighth and 10th graders constituted 44.91% (20,011/43,994) and 45.99% (19,549/43,994) of the study sample, respectively. The questions pertaining to social media use and health behaviors were administered to approximately 20% of the 12th-grade students during 2014 and 2015 compared with 100% of the 8th- and 10th-grade students. Hence, 12th-grade students constituted a smaller portion of the study sample. We limited the selection of data to 2014 and 2015 due to major changes to the questions on social media use in 2014 [24,25].

Dependent and Independent Variables

Our dependent variables were physical activity and sleep adequacy. The MTF survey measured responses to questions on physical activity ("How often do you exercise vigorously?") and sleeping habits ("How often do you get at least seven hours of sleep?") using a Likert-like scale with 5 possible responses:

XSL•FO
RenderX

never, seldom, sometimes, most days, nearly every day, and every day. We created our dependent variables by combining adjacent categories of responses. We refer to the 3 levels of outcome variables as never/seldom, sometimes, and every day/nearly every day of physical activity or daily 7 hours of sleep.

We measured the main independent variable, frequency of social media use, by how often the student used social media ("How often do you visit social networking websites like Facebook, Twitter, Instagram, etc.?") using a Likert-like scale with 5 possible responses: never, a few times a year, one to two times a month, once a week, or every day. We included the following individual level covariates in all our models: sex, average letter grade, number of hours of homework per week, grade level, and level of happiness on the day of the survey (very happy, pretty happy, or not too happy). We included school location (in a metropolitan statistical area as defined for the US Census or not) and the Census region in which the school was located to account for regional variations. We included survey year to account for differences between the survey years.

Statistical Analysis

We used a repeated cross-sectional analysis by pooling survey data from the years 2014 and 2015 for 8th-, 10th-, and 12th-grade students. We employed multivariable generalized ordered logistic regressions to compare the association of social media use with different levels of physical activity and sleep adequacy. We used generalized ordered logistic models instead of ordered logistic regression models because the ordered logistic models did not meet the proportional odds assumption [27,28]. We tested for the joint significance of different levels of social media use using the adjusted Wald test. We estimated marginal effects (MEs) for the main independent variable, social media use frequency, by holding all other variables at their observed values. All analyses used pooled survey weights derived from the sampling frame sizes provided by Inter-university Consortium for Political and Social Research, along with the individual sample weights provided in the public-use data files. We conducted all analyses using Stata 14 (StataCorp LLC).

Results

Sample Characteristics

The study sample consisted of students who answered questions on social media use and questions on one or both of the health behaviors (N=43,994). Table 1 presents the characteristics of the study sample over different levels of social media use.

As Table 1 shows, the study population consisted of 51.13% (21,276/42,067) female students; 37.48% (17,160/43,994) were from the South and 80.07% (34,953/43,994) were from a metropolitan statistical area. Approximately half of the study population reported 1-4 hours of homework per week (21,029/42,332, 50.08%) and 75.25% (33,392/42,133) reported letter grades of B– or better. Most (28,381/43,460, 65.52%) of the students reported being pretty happy. Additionally 9.12% (4434/43,994) of the students were 12th graders, and 8th and 10th graders accounted equally for the rest of the study population.

Of the 20,879 students who answered both the questions on physical activity and on social media use, almost half (n=10,033) reported being physically active every day or nearly every day, while 17.06% (n=3563) reported lower levels of physical activity (Table 1). In all levels of social media use, a higher percentage of students reported being physically active regularly and a smaller percentage reported being sedentary.

Of the 20,950 students who answered both the questions on sleep and on social media use, about 41.82% (n=8689) reported getting 7 hours of sleep nearly every day, while 18.39% (n=3986) of the students reported never or seldom getting 7 hours of sleep (Table 1). In all levels of social media use except use every day, a higher percentage of students reported getting adequate sleep. Among those who used social media every day, an equal percentage of students reported getting moderately adequate ("sometimes") or adequate levels ("every day/nearly every day") of sleep.

XSL•FO
RenderX

Table 1. Characteristics of students by frequency of social media use.

Characteristics	Social media frequency observations, n (%)						*P* value
	Never (n=3087)	A few times a year (n=990)	1-2 times a month (n=1607)	Once a week (n=4479)	Every day (n=33,831)	Total (N=43,994)	
Frequency of physical activity							<.001
Never/seldom	312 (22.3)	101 (21.0)	145 (19.0)	321 (16.7)	2684 (16.56)	3563 (17.14)	
Sometimes	499 (36.8)	160 (37)	304 (40.5)	722 (36.1)	5598 (34.36)	7283 (34.98)	
Every day/nearly every day	543 (40.9)	185 (42.0)	307 (40.5)	960 (47.2)	8038 (49.08)	10,033 (47.88)	
Frequency of sleeping 7 hours/day							<.001
Never/seldom	206 (13.8)	87 (20.0)	149 (19.2)	352 (16.3)	3192 (18.96)	3986 (18.39)	
Sometimes	449 (33.1)	162 (36.0)	270 (34.5)	763 (38.7)	6631 (40.83)	8275 (39.79)	
Every day/nearly every day	703 (53.1)	196 (44.0)	330 (46.3)	894 (45.0)	6566 (40.22)	8689 (41.82)	
Survey year							<.001
2014	1553 (50.00)	527 (53.0)	815 (49.7)	2256 (50.44)	15,935 (47.08)	21,086 (47.86)	
2015	1534 (50.00)	463 (47.0)	792 (50.3)	2223 (49.56)	17,896 (52.92)	22,908 (52.14)	
Grade							<.001
8th	1811 (58.68)	511 (50.0)	754 (45.5)	2282 (50.79)	14,653 (42.68)	20,011 (44.91)	
10th	1105 (36.56)	393 (42.4)	692 (45.3)	1780 (40.74)	15,579 (47.69)	19,549 (45.98)	
12th	171 (4.7)	86 (8.0)	161 (9.0)	417 (8.5)	3599 (9.63)	4434 (9.11)	
Sex							<.001
Male	1989 (67.46)	670 (70.7)	1066 (70.13)	2786 (63.96)	14,280 (43.53)	20,791 (48.87)	
Female	950 (32.5)	272 (29.3)	453 (29.9)	1454 (36.04)	18,147 (56.47)	21,276 (51.13)	
Region							<.001
North east	572 (16.2)	178 (17.1)	279 (15.0)	829 (17.0)	6631 (18.18)	8489 (17.78)	
North central	604 (21.3)	204 (21.4)	311 (22.1)	912 (22.1)	6812 (22.00)	8843 (21.95)	
South	1153 (35.8)	363 (35.1)	630 (36.9)	1657 (35.32)	13,357 (38.02)	17,160 (37.48)	
West	758 (26.7)	245 (26.1)	387 (26.0)	1081 (25.61)	7031 (21.80)	9502 (22.79)	
Metropolitan statistical area							.01
No	597 (19.5)	222 (22.0)	364 (22.9)	995 (21.3)	6863 (19.59)	9041 (19.93)	
Yes	2490 (80.5)	768 (78.0)	1243 (77.11)	3484 (78.74)	26,968 (80.41)	34,953 (80.07)	
Level of happiness							.03
Not too happy	502 (15.9)	159 (17)	285 (17.5)	650 (14.8)	5448 (16.34)	7044 (16.19)	
Pretty happy	1918 (63.5)	634 (65.7)	1045 (66.14)	2945 (66.75)	21,839 (65.50)	28,381 (65.52)	
Very happy	623 (20.6)	177 (17.8)	257 (16.4)	815 (18.5)	6163 (18.16)	8035 (18.29)	
Homework (hours/week)							.09
0	210 (7.3)	80 (9)	142 (8.7)	325 (8.09)	2693 (7.98)	3450 (7.99)	
1-4	1397 (47.71)	462 (49.8)	773 (50.7)	2146 (51.08)	16,521 (50.14)	21,029 (50.08)	
5-9	596 (20.1)	191 (20)	294 (19.3)	893 (20.6)	6362 (19.58)	8336 (19.72)	
10-14	337 (11.2)	104 (10.9)	152 (10.1)	422 (9.5)	3327 (9.94)	4342 (9.99)	
15-19	169 (5.6)	37 (4)	69 (5)	235 (5.1)	1781 (5.47)	2291 (5.39)	
20-24	123 (4.7)	32 (4)	43 (3)	121 (2.8)	1174 (3.50)	1493 (3.49)	
≥25	99 (3)	30 (3)	47 (3)	125 (2.8)	1090 (3.40)	1391 (3.34)	
Letter grade							<.001
D	82 (3)	33 (4)	58 (4)	102 (2.5)	856 (2.5)	1131 (2.63)	

XSL•FO
RenderX

Characteristics	Social media frequency observations, n (%)						*P* value
	Never (n=3087)	A few times a year (n=990)	1-2 times a month (n=1607)	Once a week (n=4479)	Every day (n=33,831)	Total (N=43,994)	
C–	89 (3)	32 (4)	68 (4)	147 (3.7)	1194 (3.69)	1530 (3.68)	
C	117 (4.4)	48 (5)	92 (6)	278 (6.6)	1796 (5.55)	2331 (5.59)	
C+	173 (5.6)	81 (9)	139 (9.3)	403 (9.9)	2953 (9.09)	3749 (8.93)	
B–	203 (6.4)	102 (10.4)	160 (10.4)	455 (10.7)	3308 (10.33)	4228 (10.11)	
B	374 (13.3)	148 (14.4)	236 (15.7)	698 (16.5)	4968 (15.16)	6424 (15.16)	
B+	461 (15.4)	161 (17.3)	255 (18.5)	676 (15.6)	5725 (17.71)	7278 (17.35)	
A–	592 (20.6)	156 (18.4)	231 (15.7)	724 (17.2)	5707 (17.49)	7410 (17.62)	
A	835 (28.4)	177 (18.8)	256 (16.3)	771 (17.3)	6013 (18.42)	8052 (18.93)	

Associations Between Dependent and Independent Variables

Results of multivariable regressions for physical activity indicated a statistically significant association between frequencies of social media use and physical activity (adjusted Wald test $F_{8,20871}$=18.39; P<.001); and between frequencies of social media use and sleep adequacy (adjusted Wald test $F_{8,20942}$=2.05; P=.04). Table 2 shows the MEs of different impacts of social media use on different levels of physical activity and sleep adequacy. We calculated the MEs holding all other covariates at their observed values.

Among students with a low level of physical activity, the likelihood of reporting physical activity decreased with increasing social media use (for social media nonusers, ME 26.4%, 95% CI 23.4%-29.4%, and for daily social media users, ME 15.8%, 95% CI 15.1%-16.4%). Among students with a higher level of physical activity, the likelihood of reporting physical activity increased with increasing frequency of social media use (for social media nonusers, ME 35.5%, 95% CI 32.6%-38.5%, and for daily social media users, ME 50.1%, 95% CI 49.2%-51.0%). In contrast to this linear relationship, the relationship among moderately active students peaked with students who used social media once or twice a month reporting the highest likelihood of vigorous exercise (ME 42.0%, 95% CI 37.6%-46.3%).

Among students who got inadequate sleep, social media nonusers were less likely to report getting 7 hours of sleep regularly (ME 16.6%, 95% CI 14.2%-18.9%) than were students who used social media every day (ME 18.3%, 95% CI 17.6%-19.0%). Among those who got adequate sleep, social media nonusers were more likely to report getting 7 hours of sleep (ME 46.6%, 95% CI 43.4%-49.7%) than were students who were daily social media users (ME 41.3%, 95% CI 40.4%-42.1%). Students who reported moderately adequate sleep and who used social media once or twice a month were least likely report adequate sleep (ME 36.4%, 95% CI 32.2%-40.6%).

Additional analysis including parental education and race/ethnicity as control variables did not significantly change the reported results. However, including these variables resulted in a smaller sample size, as one-third of the students did not provide parental education data and almost 29.92% (13,165/43,994) of the observations lacked data on race/ethnicity. Hence, we report the original results.

Table 2. Regression results showing the association between health behaviors and frequency of social media use by middle and high school students[a,b,c].

Social media use frequency	% marginal effect (95% CI)					
	Vigorous physical activity (n=19,543)			Sleeping 7 hours/day (n=19,596)		
	Never/seldom	Sometimes	Every day	Never/seldom	Sometimes	Every day
Never	26.4 (23.4-29.4)	38.1 (34.8-41.3)	35.5 (32.6-38.5)	16.6 (14.2-18.9)	36.9 (33.7-40.0)	46.6 (43.4-49.7)
A few times a year	24.4 (19.8-28.9)	37.8 (32.4-43.2)	37.8 (32.5-43.1)	21.4 (16.9-25.9)	37.9 (32.6-43.3)	40.7 (35.4-46.0)
1-2 times a month	20.4 (16.9-24.0)	42.0 (37.6-46.3)	37.6 (33.6-41,6)	19.5 (16.2-22.9)	36.4 (32.2-40.6)	44.0 (39.9-48.2)
Once a week	18.6 (16.4-20.8)	36.5 (33.9-39.1)	44.9 (42.3-47.5)	17.4 (15.4-19.3)	40.1 (37.5-42.8)	42.5 (40.0-45.0)
Every day	15.8 (15.1-16.4)	34.1 (33.2-35.0)	50.1 (49.2-51.0)	18.3 (17.6-19.0)	40.4 (39.5-41.4)	41.3 (40.4-42.1)

[a]Generalized ordered logistic regression models are used for the estimation. Marginal effect indicates predicted percentage of students with different intensities of health behaviors and social media use, holding all the other covariates at their observed values.

[b]All results are statistically significant at P<.001.

[c]Data are from the Monitoring the Future survey (2014 and 2015) with students from 8th, 10th, and 12th grades.

XSL•FO
RenderX

Discussion

Principal Findings

Our study results indicate that the association between health behaviors and the frequency social media use is significant and the direction of the association differs by whether the health behaviors are at high, medium, or low levels. The health behaviors at both the high and the lower end of the spectrum are reinforced further with more frequent social media use, whereas moderate health behaviors show an interesting nonlinear relationship with the frequency of social media use. We adjusted for reported happiness to account for stress, which may adversely influence health behaviors. While having a higher homework load may take time away from physical activity, sleep, and social media use, we controlled for the homework hours in all models. We also adjusted for grade level, as students in higher grades may have less time for sleep and physical activities due to more academic demands on their time.

The results show that students who were moderately active (7283/20,879, 34.98% of the study sample) or who reported moderately adequate sleep (8275/20,950, 39.79% of the study sample) had a nonlinear relationship with the frequency of social media use. This relationship indicates that a moderate amount of social media use may have a beneficial impact on moderate physical activity and a detrimental influence on moderately adequate sleep among students. Most of the previous research did not delineate different levels of health behaviors, leading to the oversimplified conclusion that more social media use is associated with worse health behaviors. Our results are in line with the emerging literature on the relationship between social media use intensity and other health outcomes, including a U-shaped, nonlinear "Goldilocks effect" of the intensity of internet use on depression and overall mental well-being among adolescents [21,23]. The Goldilocks effect refers to the phenomenon where a moderate intensity of social media use is associated with the highest levels of mental well-being, compared with either too much or too little social media use.

These results are in line with the evidence that users extend their offline personality and behaviors into their online activities [29]. Our results indicate that, among physically active students, frequent social media use was associated with a higher likelihood of daily vigorous physical activity, which may be due to a positive feedback loop from sharing related news or pictures on social media. On the other hand, sedentary behaviors changed for the worse with increased social media use, which may further the supposition that time spent on social media is not available for other activities. Sleeping behaviors of both inadequate and adequate sleepers were reinforced with more frequent social media use. Inadequate sleepers may have had difficulty falling asleep or may have been light sleepers, and such issues were magnified with more social media use. On the other hand, adequate sleepers may have developed sleeping habits that were immune to social media use.

Implications

These results have several implications for adolescents as their social media use has drastically increased over time. Adolescents' social interactions have increasingly moved to social media platforms. To capitalize on this trend, there has been a marked shift in promoting health behaviors and providing interventions using social media platforms, which contribute to further increases in the frequency of social media use. While there is cause for concern, our study indicates some opportunities. There may be a "sweet spot" where a moderate level of social media use may provide maximum or minimum benefit for a particular health behavior for a specific set of students. The results also suggest that a one-type-fits-all approach may be detrimental to those who are already in the group that is at most risk of adverse behaviors.

While students may engage in vigorous physical activity due to participation in school teams and other organized activities, the levels of physical activity tend to decrease as they graduate from high school, go to college where activities are less structured, and eventually settle into adulthood, when they have to be self-motivated to remain physically active [30]. Unlike physical activity, sleep patterns tend to fluctuate more; however, most of the adult population falls into a moderately adequate sleep category with average reported sleep of less than 6 hours per day [31]. Further, average use of social media does not seem to decrease much as teenagers settle into early adulthood [4,32]. While the relationship between healthy behaviors and social media use in late adulthood may be influenced by additional factors, it is important to understand this nonlinear relationship during early adulthood, when many of the health behaviors are formed. Social media use is shaping up to be a large influencer of health behaviors, with social media-based programs being developed to improve fitness, reduce stress, share experiences on health behaviors, and form support groups to improve health and wellness for all age groups [7,33]. However, there is also growing evidence that social media may promote negative emotions due to excessive social comparison, including those on health behaviors [34-36]. Hence, it is imperative to understand whether there is an optimal level of social media use that has the best impact on health behaviors.

Limitations

This was a cross-sectional study and, hence, we cannot claim to assess causality between health behaviors and social media use. The study population was a nationally representative sample of students. However, caution should be exercised when generalizing the results to other adolescents who are home schooled or who have dropped out of school.

While there was a survey question on the number of hours spent on social media, a large percentage of the answers were either missing or inconsistent and, hence, not used in this study. Since time used for social media was affecting health behaviors [6,11,13], future studies would benefit from including the actual time spent on social media activities rather than categories.

The data did not include information on specific social media sites visited or on specific activities conducted on social media (such as posting and responding to messages, or creating content). While results indicate reinforcement of both high and low levels of health behaviors, understanding the content and its specific methods of use by students would provide deeper insights into understanding the nonlinear relationship among

XSL•FO
RenderX

those with moderate health behaviors. Inadequate sleep may be worsened further by social media content.

Students who exercise vigorously may be involved in organized sports activities such as school athletic programs or competitive sports outside the school and, hence, may be more disciplined in their commitment to physical activities. In that case, the effect of social media use on their physical activity levels may be overstated. We controlled for this to some extent by adjusting for academic achievement as a proxy for their intrinsic motivation. Students who sleep less may do so due to issues related to anxiety, depression, or other stressors, which may also influence use of social media. We accounted for this by including level of happiness as a proxy for mental well-being. The survey question on sleep asked about the regularity of getting 7 hours of sleep, whereas the US National Sleep Foundation and American Academy of Sleep Medicine recommend 8 to 10 hours of sleep for teenagers [37]. Hence, the survey possibly underrepresented the inadequacy of sleep. Future studies should include variables that capture sleep quality or problems in addition to questions on sleep duration.

Conclusion

Our study indicates that regular social media use reinforces health behaviors of adolescents at the extreme ends of the health behavior spectrum. Among adolescents who follow health behaviors to a moderate extent, a moderate use of social media provides the most benefit. Future studies examining the content of social media sites visited and the interaction patterns, as well as the number of hours spent on social media sites, would inform our understanding of the impact of social media use on health behaviors and facilitate using optimal social media to promote health behaviors.

Acknowledgments

This work was supported by the National Institute of General Medical Sciences of the US National Institutes of Health under award numbers UL1GM118979, TL4GM118980, and RL5GM118978 and by a Summer Research Assistantship from the Office of Research and Sponsored Programs at California State University, Long Beach, Long Beach, CA, USA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Villanti AC, Johnson AL, Ilakkuvan V, Jacobs MA, Graham AL, Rath JM. Social media use and access to digital technology in US young adults in 2016. J Med Internet Res 2017 Jun 07;19(6):e196 [FREE Full text] [doi: 10.2196/jmir.7303] [Medline: 28592394]
2. Children and parents: media use and attitudes report. London, UK: Ofcom; 2017 Nov 29. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0020/108182/children-parents-media-use-attitudes-2017.pdf [accessed 2019-06-06] [WebCite Cache ID 78w0ojT8S]
3. Duggan M, Ellison NB, Lampe C, Lenhart A, Madden M. Social media update 2014. Washington, DC: Pew Research Center; 2015 Jan 09. URL: https://www.pewresearch.org/wp-content/uploads/sites/9/2015/01/PI_SocialMediaUpdate20144.pdf [accessed 2019-06-06] [WebCite Cache ID 78w14Jy9d]
4. Lenhart A, Purcell K, Smith A, Zickuhr K. Social media & mobile internet use among teens and young adults. Washington, DC: Pew Internet & American Life Project; 2010 Feb 03. URL: https://www.pewinternet.org/2010/02/03/social-media-and-young-adults/ [accessed 2019-06-06]
5. Jelenchick LA, Eickhoff J, Christakis DA, Brown RL, Zhang C, Benson M, et al. The Problematic and Risky Internet Use Screening Scale (PRIUSS) for Adolescents and Young Adults: scale development and refinement. Comput Human Behav 2014 Jun 01;35:171-178 [FREE Full text] [doi: 10.1016/j.chb.2014.01.035] [Medline: 24882938]
6. Bruni O, Sette S, Fontanesi L, Baiocco R, Laghi F, Baumgartner E. Technology use and sleep quality in preadolescence and adolescence. J Clin Sleep Med 2015 Dec 15;11(12):1433-1441 [FREE Full text] [doi: 10.5664/jcsm.5282] [Medline: 26235161]
7. Vaterlaus JM, Patten EV, Roche C, Young JA. #Gettinghealthy: the perceived influence of social media on young adult health behaviors. Comput Hum Behav 2015 Apr;45:151-157. [doi: 10.1016/j.chb.2014.12.013]
8. Council on Communications and Media. Media use in school-aged children and adolescents. Pediatrics 2016 Nov;138(5):e20162592. [doi: 10.1542/peds.2016-2592] [Medline: 27940794]
9. Booker CL, Skew AJ, Kelly YJ, Sacker A. Media use, sports participation, and well-being in adolescence: cross-sectional findings from the UK household longitudinal study. Am J Public Health 2015 Jan;105(1):173-179. [doi: 10.2105/AJPH.2013.301783] [Medline: 25494209]
10. Börnhorst C, Wijnhoven TMA, Kunešová M, Yngve A, Rito AI, Lissner L, et al. WHO European Childhood Obesity Surveillance Initiative: associations between sleep duration, screen time and food consumption frequencies. BMC Public Health 2015 Apr 30;15:442 [FREE Full text] [doi: 10.1186/s12889-015-1793-3] [Medline: 25924872]

XSL•FO
RenderX

11. Arora T, Broglia E, Thomas GN, Taheri S. Associations between specific technologies and adolescent sleep quantity, sleep quality, and parasomnias. Sleep Med 2014 Feb;15(2):240-247. [doi: 10.1016/j.sleep.2013.08.799] [Medline: 24394730]
12. Fobian AD, Avis K, Schwebel DC. Impact of media use on adolescent sleep efficiency. J Dev Behav Pediatr 2016 Jan;37(1):9-14 [FREE Full text] [doi: 10.1097/DBP.0000000000000239] [Medline: 26651090]
13. Reynolds AC, Meltzer LJ, Dorrian J, Centofanti SA, Biggs SN. Impact of high-frequency email and instant messaging (E/IM) interactions during the hour before bed on self-reported sleep duration and sufficiency in female Australian children and adolescents. Sleep Health 2019 Dec;5(1):64-67. [doi: 10.1016/j.sleh.2018.10.008] [Medline: 30670168]
14. Levenson JC, Shensa A, Sidani JE, Colditz JB, Primack BA. The association between social media use and sleep disturbance among young adults. Prev Med 2016 Apr;85:36-41 [FREE Full text] [doi: 10.1016/j.ypmed.2016.01.001] [Medline: 26791323]
15. Woods HC, Scott H. #Sleepyteens: social media use in adolescence is associated with poor sleep quality, anxiety, depression and low self-esteem. J Adolesc 2016 Dec;51:41-49. [doi: 10.1016/j.adolescence.2016.05.008] [Medline: 27294324]
16. Xanidis N, Brignell CM. The association between the use of social network sites, sleep quality and cognitive function during the day. Comput Hum Behav 2016 Feb;55:121-126. [doi: 10.1016/j.chb.2015.09.004]
17. Dijk DJ. Not so fast: sleep and media use. J Sleep Res 2014 Aug;23(4):363 [FREE Full text] [doi: 10.1111/jsr.12212] [Medline: 25073803]
18. Tavernier R, Willoughby T. Sleep problems: predictor or outcome of media use among emerging adults at university? J Sleep Res 2014 Aug;23(4):389-396. [doi: 10.1111/jsr.12132] [Medline: 24552437]
19. Meena PS, Mittal PK, Solanki RK. Problematic use of social networking sites among urban school going teenagers. Ind Psychiatry J 2012 Jul;21(2):94-97. [doi: 10.4103/0972-6748.119589] [Medline: 24250039]
20. Moreno MA, Jelenchick L, Cox E, Young H, Christakis DA. Problematic internet use among US youth: a systematic review. Arch Pediatr Adolesc Med 2011 Sep;165(9):797-805 [FREE Full text] [doi: 10.1001/archpediatrics.2011.58] [Medline: 21536950]
21. Bélanger RE, Akre C, Berchtold A, Michaud P. A U-shaped association between intensity of internet use and adolescent health. Pediatrics 2011 Feb;127(2):e330-e335. [doi: 10.1542/peds.2010-1235] [Medline: 21242218]
22. Li W, O'Brien JE, Snyder SM, Howard MO. Characteristics of internet addiction/pathological internet use in U.S. university students: a qualitative-method investigation. PLoS One 2015;10(2):e0117372 [FREE Full text] [doi: 10.1371/journal.pone.0117372] [Medline: 25647224]
23. Przybylski AK, Weinstein N. A large-scale test of the Goldilocks hypothesis. Psychol Sci 2017 Dec;28(2):204-215. [doi: 10.1177/0956797616678438] [Medline: 28085574]
24. Johnston LD, Bachman JG, O'Malley PM, Schulenberg JE, Miech RA. Monitoring the Future: A Continuing Study of American Youth (8th- and 10th-Grade Surveys). Ann Arbor, MI: Inter-university Consortium for Political and Social Research; 2016.
25. Johnston LD, Bachman JG, O'Malley PM, Schulenberg JE, Miech RA. Monitoring the Future: A Continuing Study of American Youth (12th-Grade Survey). Ann Arbor, MI: Inter-university Consortium for Political and Social Research; 2017.
26. Bachman JG, Johnston LD, O'Malley PM, Schulenberg JE, Miech RA. The Monitoring the Future Project after Four Decades: Design and Procedures. Monitoring the Future Occasional Paper Series. Paper 82. Ann Arbor, MI: Institute for Social Research, University of Michigan; 2015.
27. Long SJ, Long JS, Freese J. Regression Models for Categorical Dependent variables Using Stata. College Station, TX: Stata Press; 2006.
28. Williams R. Understanding and interpreting generalized ordered logit models. J Math Sociol 2016 Jan 29;40(1):7-20. [doi: 10.1080/0022250X.2015.1112384]
29. Gosling SD, Augustine AA, Vazire S, Holtzman N, Gaddis S. Manifestations of personality in online social networks: self-reported Facebook-related behaviors and observable profile information. Cyberpsychol Behav Soc Netw 2011 Sep;14(9):483-488 [FREE Full text] [doi: 10.1089/cyber.2010.0087] [Medline: 21254929]
30. Gordon-Larsen P, Nelson MC, Popkin BM. Longitudinal physical activity and sedentary behavior trends: adolescence to adulthood. Am J Prev Med 2004 Nov;27(4):277-283. [doi: 10.1016/j.amepre.2004.07.006] [Medline: 15488356]
31. Ford ES, Cunningham TJ, Croft JB. Trends in self-reported sleep duration among US adults from 1985 to 2012. Sleep 2015 May 01;38(5):829-832 [FREE Full text] [doi: 10.5665/sleep.4684] [Medline: 25669182]
32. Perrin A. Social networking usage: 2005-2015. Washington, DC: Pew Research Center; 2015 Oct 08. URL: https://www.pewresearch.org/wp-content/uploads/sites/9/2015/10/PI_2015-10-08_Social-Networking-Usage-2005-2015_FINAL.pdf [accessed 2019-06-06] [WebCite Cache ID 78w1Yz4qH]
33. Looyestyn J, Kernot J, Boshoff K, Maher C. A web-based, social networking beginners' running intervention for adults aged 18 to 50 years delivered via a Facebook group: randomized controlled trial. J Med Internet Res 2018 Feb 26;20(2):e67 [FREE Full text] [doi: 10.2196/jmir.7862] [Medline: 29483065]
34. Tiggemann M, Slater A. NetGirls: the internet, Facebook, and body image concern in adolescent girls. Int J Eat Disord 2013 Sep;46(6):630-633. [doi: 10.1002/eat.22141] [Medline: 23712456]
35. Lee SY. How do people compare themselves with others on social network sites? The case of Facebook. Comput Hum Behav 2014 Mar;32:253-260. [doi: 10.1016/j.chb.2013.12.009]

XSL•FO
RenderX

36. Haferkamp N, Krämer NC. Social comparison 2.0: examining the effects of online profiles on social-networking sites. Cyberpsychol Behav Soc Netw 2011 May;14(5):309-314. [doi: 10.1089/cyber.2010.0120] [Medline: 21117976]
37. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's updated sleep duration recommendations: final report. Sleep Health 2015 Dec;1(4):233-243. [doi: 10.1016/j.sleh.2015.10.004] [Medline: 29073398]

Abbreviations

MTF: Monitoring the Future
ME: marginal effect

Edited by G Eysenbach; submitted 05.04.19; peer-reviewed by M El Tantawi, N Aly; comments to author 22.05.19; revised version received 23.05.19; accepted 23.05.19; published 18.06.19.

Please cite as:
Shimoga SV, Erlyana E, Rebello V
Associations of Social Media Use With Physical Activity and Sleep Adequacy Among Adolescents: Cross-Sectional Survey
J Med Internet Res 2019;21(6):e14290
URL: http://www.jmir.org/2019/6/e14290/
doi:10.2196/14290
PMID:31215512

©Sandhya V Shimoga, Erlyana Erlyana, Vida Rebello. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Not Just a Headache: Qualitative Study About Web-Based Self-Presentation and Social Media Use by People With Migraine

Carly Pearson[1], BSc, MSc; Rosanna Swindale[1], BSc; Peter Keighley[1], BSc; Alison Ruth McKinlay[1], BA (Hons), PhD; Leone Ridsdale[1], MD, PhD

King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, United Kingdom

Corresponding Author:
Leone Ridsdale, MD, PhD
King's College London
Institute of Psychiatry, Psychology and Neuroscience
16 De Crespingy Park
PO Box 63
London, SE5 8AF
United Kingdom
Phone: 44 02078480815
Email: leone.ridsdale@kcl.ac.uk

Abstract

Background: To help with a long-term but invisible medical condition such as migraine, many people seek information and support on social media. The effect of using social media for people with migraine is not fully understood and remains to be investigated.

Objective: The aim of this study was to describe how people with migraine use social media and how social media use affects their identity and sense of self.

Methods: A total of 20 participants who experienced migraine were recruited via migraine-specific charities. Semistructured interviews were conducted with questions based on a topic guide. Interviews were transcribed verbatim, and transcripts were analyzed using thematic analysis.

Results: People with migraine are using social media to obtain information to better understand their condition and treatment options. Social media offers instant access to continuous information and social support. This exchange of social support and information was viewed as mutually beneficial. Participants viewed social media as an outlet to vent frustrations and validate the migraine experience. Several participants pointed out that the invisible and episodic nature of migraine can lead to societal misunderstanding of the impact and or severity of their condition. Some participants masked their online migraine-related behavior using different sites or closed online groups to control who saw their migraine-related content. Participating in closed social media groups sometimes changed Web-based behavior in other areas of the platform. This illustrates the complex relationship between migraine, social media, and identity.

Conclusions: How migraine is part of an individual's identity and how this is represented online can vary. Social media can provide people who experience migraine with instant and continuous access to support and information, from a group of empathic others with similar lived experiences. Social media is used to validate the illness experience, as well as provide reassurance and help reduce feelings of isolation.

(J Med Internet Res 2019;21(6):e10479) doi:10.2196/10479

KEYWORDS

migraine; internet; social media; eHealth; social support; self-management; qualitative research

Introduction

Migraines are severe headaches that affect at least 5% of men and 15% of women [1]. Symptoms can include heightened sensitivity to external stimuli, nausea, vomiting, and sensory disturbances (auras) [1,2]. Between attacks, people with migraine display no physical symptoms and might choose to conceal their condition, known as *passing* [3-5]. Successful passing allows those with invisible illnesses to present themselves in ways that are not defined by their impairment

XSL•FO
RenderX

[3]. Despite this, migraine is a condition that can affect identity and self-concept. Episodes can place burden directly on the individual but also have indirect effects on wider society. It is now the commonest cause of disability in people of working age [6].

Migraine symptoms are typically episodic and unpredictable, making it a difficult condition to manage in daily life. It is distinctive from other chronic conditions as it lacks a visible illness marker [7]. Kleinman proposed that such problems can cause frustration from not being believed or understood and result in altered self-concept or self-esteem [8]. People with migraine may not receive adequate support owing to the experience of illness and pain being subjective [7]. Others might not believe or understand the impact of migraine symptoms [9]. This group may therefore feel dissatisfied with their support and medical treatment [10-12], which can make the experience of receiving validation and acknowledgement from others especially valuable.

Living with a long-term condition requires self-management, where individuals take control of their health care through targeted symptom management and lifestyle changes [12]. Internet-based technologies offer a way to self-manage, and more recently, social media has enabled the rapid exchange of information [13] for this purpose. Social media includes internet-based websites and mobile phone apps that enable users to generate content and interact with others [14,15]. The growth of social media has meant that people with health conditions can rapidly acquire and share health-related information [16]. Researchers have called for further study into the precursors of lasting health-related change through the use of social media [17].

Social identity theory could explain some behavior change following Web-based interactions. If behaviors are linked to the role a person occupies within a social group (ie, son, parent, and teacher) [18], and social norms tied to group membership [19], this may apply to online group interactions as well. Under social identity theory [20], once an individual identifies with a social group, they may strive for a self-concept [18,21] associated with their group membership. In this study, we elected to focus on identity and presentation of *the self* among people with migraine in Web-based spaces. For the purposes of this discussion, the self is conceptualized as an accumulation of the personas that are represented in a broad set of social interactions [22].

Social media use has been reported to increase self-management in other long-term conditions such as diabetes [23-26]. However, there are likely to be differences in the perceived affordances of social media among medical conditions, and online self-presentation in migraine has yet to be investigated. Interest in the use of social media in migraine and how individuals perceive it is growing [13,27]. The principal investigator in this study is a neurologist who runs a headache clinic and evaluates health interventions designed to improve self-management using qualitative methods [28,29]. Funding from the European Research Council allowed the group to extend their research, aiming to explore how people with migraine independently use social media in the context of their lives and how this use impacts them.

Further exploration of behavior change following social media use in people with chronic long-term conditions is warranted [17]. One possible mechanism is that users must curate their identity and self-presentation using social media platforms. Therefore, we asked the following: What are the uses of social media among people with migraines? How does use of social media affect an individual's sense of self and online identity?

Methods

Study Design

We used qualitative methodology aiming to gather in-depth views and experiences of using social media in the context of migraines. This approach permits open exploration of participants' views, and the interviewer can adapt questioning based on each participant's response. Ethical approval was given by the Psychiatry, Nursing and Midwifery Research Ethics Subcommittees, King's College London (reference: PNM/13/14-18).

Participants and Recruitment

Participants expressed interest in response to advertisements placed on social media sites and newsletters of the charities: Migraine Trust and Migraine Action. Inclusion criteria were adults aged ≥18 years (no upper age limit), self-report experience of migraines, use social media in the context of migraine, fluent in English, and live within a catchment area of London and Birmingham in the United Kingdom.

A total of 35 eligible participants responded from all over the United Kingdom. Participants were selected depending on their proximity to the place of work of the interviewers, which was London and Birmingham. They were then invited for an interview in their home or a public space near their homes. In total, 2 participants selected for the interview did not participate, one owing to work commitments and another was uncontactable. All participants received an information sheet explaining the aims and methods of the study and gave written informed consent to participate. Recruitment ended at 20 participants when data saturation was reached.

Data Collection

One-to-one semistructured interviews were conducted by the researchers who were trained in qualitative methods with the experience of working with patients and research participants. The interviews lasted approximately 30 to 45 min. Interviews took place between November 2016 and March 2017. An interview schedule was devised with questions developed from the existing literature to explore the study aims (Table 1). Pilot interviews (P5, P10, and P19) were conducted initially to assess the viability of the interview schedule, and minor iterations were implemented to ensure clarity of the questions. This involved clarifying language in the question around identity and subsequent probes.

XSL•FO
RenderX

Table 1. Interview topics and probes.

Questions	Probes
When did you first start to experience migraine?	How did you feel when you first started to experience migraine?
Could you tell me what personal research have you done around migraine? (eg, family/friends, self-help books)	Do you know anyone else in real life who experiences migraine?
Could you talk about any social media platforms you have used specifically relating to migraine? (eg, Twitter, Facebook, blogging)	What were your motivations for you to visit this website or use this app?
Have you ever found information or support for your migraines on social media that you didn't find elsewhere?	Has being part of the online migraine community impacted on your migraine management?
Could you describe how you use social media in the context of your migraine to find out new information and advice?	Do you use the official NHS[a] choices website to gather information about migraine?
Could you describe how you use social media in the context of your migraine to share information and advice?	Would you say you primarily seek to find out information or to share it?
Have you found communicating online with others who have migraine a supportive experience?	Is the social support online different to offline?
Could you describe any benefits of having an online presence?	What can social media offer that other sources of information cannot?
Could you describe any drawbacks of having an online presence?	Do you find the information to be accurate and reliable?
Do you see social media platforms as a way of gaining expertise in migraine knowledge?	How does this compare to traditional methods of speaking one-to-one in person with a healthcare professional?
What kind of impact has using social media had on your identity?	How do your migraines fit into your online identity?
Have you ever had a bad reaction from other people in relation to experiencing migraines?	Has social media helped you to deal with this reaction?
Is there any way you would like to see social media platforms for migraine improved?	Potential areas: Convenience, quality, comprehensiveness, immediacy of information, user demographics, user uptake, privacy.
Would you recommend that other people who experience migraine use social media?	Could you explain why you would or wouldn't recommend it?

[a]NHS: National Health Service.

Data Analysis

Interviews were audio-recorded and transferred to a password protected computer. To preserve participant anonymity, identifying data were not transcribed. Audio and transcript files were identifiable only by an arbitrary participant number. Interviews were transcribed verbatim by a third-party transcription service and checked by interviewers. The transcripts were analyzed using thematic analysis which involves inductive interpretation [30,31]. No external sources other than participant transcripts were used during the analysis process. The transcripts were coded line-by-line using NVivo 11 software for qualitative data. Similar codes were grouped together, and an iterative approach was used to create subthemes. These were further grouped together into overarching themes. A discussion of the themes and interpretations was conducted by all authors. The final interpretation of codes and themes was completed by the senior researchers (CP and LR).

Results

Participant Characteristics

Of the 20 participants interviewed, 17 were women and all were white (Table 2). The age range was 24 to 59 years, mean 39 years. Almost three-quarters (70%) had graduate or postgraduate level education. A total of 11 participants were in full-time employment and 6 were unemployed. All reported a diagnosis of migraine, with the frequency ranging from 1 day per month to daily (mean 12 days per month).

Social Media Use

Facebook was the most commonly used social media platform (19 participants). Many sought information and social support from closed migraine-specific Facebook groups. Blogs were used by 2 participants. A total of 3 participants used YouTube to observe migraine symptoms and learn about migraine causes or pain reduction. One participant created YouTube videos to educate others about migraine. Twitter was often viewed as a more professional domain rather than personal. Participants' use of social media tended to fit on a spectrum from actively engaging with social media to more observational social media users, and some people in between.

XSL•FO
RenderX

Table 2. Participant demographics.

Participant	Sex	Age (years)	Highest qualification	Employment	Living situation	Migraine days per month (n)	Social media sites used in context of migraines
P1	Female	31	Postsecondary	Unemployed	With others	3	F[a], T[b], I[c]
P2	Female	52	Postsecondary	Full-time	Alone	12	F
P3	Female	28	Postgraduate	Student	With others	2	F, T, Y[d]
P4	Female	35	Postgraduate	Part-time	With others	9	F, T, Y, B[e]
P5	Female	31	Unknown	Employed	With others	8	F
P6	Male	43	Postsecondary	Full-time	With others	4	F, T
P7	Female	23	Postgraduate	Full-time	With others	16	F, I
P8	Male	24	Postgraduate	Student	With others	1-2	F, T
P9	Male	25	Postsecondary	Unemployed	With others	1-2	F, Y
P10	Female	47	Graduate	Unemployed	With others	29	F, M[f], V[g]
P11	Female	29	Postgraduate	Full-time	With others	1	F, I
P12	Female	45	Postgraduate	Full-time	With others	12	F, P[h]
P13	Female	33	Graduate	Full- time	With others	10-15	Y, I
P14	Female	32	Postgraduate	Part-time	With others	Daily	F
P15	Female	58	Postgraduate	Part-time	With others	5	F, H[i], B, MT[j], Migraine Buddy app
P16	Female	31	Graduate	Unemployed	With others	5	F
P17	Female	59	Graduate	Retired	With others	15	F, I
P18	Female	54	Postsecondary	Unemployed	With others	3-5	F
P19	Female	47	Postsecondary	Part-time	With others	15	F
P20	Female	55	Postgraduate	Unemployed	With others	Daily	F

[a]F: Facebook.

[b]T: Twitter.

[c]I: Instagram.

[d]Y: YouTube.

[e]B: blogs.

[f]M: Migraine.com (information website with discussion facilities).

[g]V: Vestibular migraine website with online community.

[h]P: Pinterest.

[i]H: Health Unlocked (online social health network).

[j]MT: My Migraine Team (a social network for people with migraine).

Theme 1: Information Exchange

Seeking Information

All participants used social media to obtain information about migraine. In total, 11 said they gained expertise in migraine knowledge and 13 gained awareness of new treatments. For some, interacting with others in Facebook groups resulted in a change of self-management for migraines. For example:

> *I follow some online Facebook groups or pages on politics and health policies. And that's really helpful...in fact it was through that, that I pursued TMS [Transcranial Magnetic Stimulation].* [P10]

> *...you can go back to your doctor, or your specialist or whatever, and say, "I haven't tried this one," or, "People have said that this works in conjunction with that," or whatever, "What do you think?"* [P18]

Sharing Information

In total, 14 participants said they shared information with other users on social media, mostly on Facebook. Some shared information with others who experience migraines in migraine-specific groups:

> *If I see an article that I think would be useful or a diagram or something which I think would be useful to other groups I would usually share it.* [P20]

XSL•FO
RenderX

Others used social media to share information with people who do not experience migraine by posting on their personal Facebook *walls*, seeking to help others to understand their condition:

> *I just bombarded all my family and friends with every article I could find...this is what's happening to me.* [P16]

Pooling and Exchanging Information and Experiences

In total, 16 participants spoke about the benefit of pooling knowledge on social media. The dialogue and collaboration with other users added another level of benefit to information seeking and sharing:

> *...what used to happen before social media, you would research it by just Googling it and then wading your way through different articles or going to different doctors. Whereas now you can be part of a group and get first-hand experience and knowledge.* [P18]

A total of 12 participants said information gathered on social media had increased their confidence, knowledge, and skills in managing their own health care:

> *It's given me information, like, because I might mention a side effect for a drug, and the doctor might be like, "Oh, I've not heard anyone having that." And I'll be like, "Well, actually, I've spoken to various people on Twitter and they all have it," so it gives me a bit more, sort of, clout with what I'm saying...Which otherwise...you can feel very, like the only one experiencing that.* [P4]

Others benefited from reading the online discussion and not taking part themselves, whereas 8 participants referred to the benefit of sharing advice:

> *...[sharing advice] gets away from the feeling that we're just an emotional support group. No, we're not, we can actually give people the information that they need to make their lives easier and that feels good.* [P20]

Theme 2: Social Support

Reducing Isolation

A total of 10 participants described not feeling alone and that social media had helped them to feel less isolated:

> *It's comforting to know that I'm not on my own.* [P2]

In total, 9 participants spoke of having to cancel plans owing to their migraines. For some, social media provided a source of support for an unpredictable and invisible illness:

> *...when you cancel on them [offline contacts] for the fifth time you can see after a while it starts grating and rubbing on them...Then when you've got like a group that you can go to...and say, "I've got a migraine coming on again," or whatever, you don't get, "Oh, you know, you let me down last week,"...You get sympathy and empathy and understanding and recommendations and help which is nice.* [P18]

Reassurance

For 19 participants, the process of being able to hear about others' experiences and compare them with their own provided a sense of comfort:

> *It is always more reassuring when their story fits your story.* [P9]

In the cases where people were unsure of what they were experiencing, reading similar accounts from others provided validation and reminded them that they were not alone. After accessing content on social media, some participants benefited from reassurance regarding *unusual* symptoms:

> *...on YouTube...People have made videos of how their...aura then kind of grows and disappears, and I've sort of watched it...to get reassurance that this isn't me, sort of, going a bit mad...other people experience similar symptoms.* [P3]

In this sense, the use of social media served to normalize what some felt might be abnormal. In total, 6 participants also described social media as a *lifeline*:

> *I don't know how people survived beforehand actually, especially because it's [migraine] invisible.* [P20]

Forming a Continuous Support Group

A total of 10 participants referred to social media being available all the time, providing a continual source of contact with other users. In total, 7 participants described how social media groups can act as a support group:

> *...you can do it anytime, it's not like a traditional support group where you'd have to go, you'd have to wait a week or a month...* [P14]

A total of 9 participants said that offering social support can give a reciprocal benefit:

> *...it's mutually supportive because you can help them which in turn makes you feel better...* [P18]

Social media was used to gain access to an empathetic audience that is not limited to geographic location:

> *...so I went on to Facebook and I put hemiplegic migraine and I found this amazing support group of people.* [P16]

Theme 3: Validation of Migraine Experience

Understanding Migraine

A total of 14 participants referred to migraine being an invisible illness, with 10 participants saying they had been given patronizing or unhelpful advice offline by others who often saw migraine as *just a headache*. In total, 18 participants discussed how the use of social media can help validate the migraine experience and combat the lack of understanding about the unpredictable and invisible nature of migraine:

> *I think the worst thing for people is not getting support...I think social media can be a good way of calling that out when we see it and people going:*

XSL•FO
RenderX

> *"Yes, that happened to me. That's not okay." There's quite a lot of validating involved.* [P20]

> *I think it's [social media] made me feel a bit better myself, really in that...It [migraine] is a genuine thing. It's not just something that you bring on yourself because you are a wimp...* [P17]

The Value of Lived Experiences of Others

In total, 16 participants discussed the dichotomy between subjective experience versus verified facts and medical knowledge. Although medical fact seeking was important, there was a desire to hear about the subjective experience of others:

> *[I look on social media]...60% for the objective stuff...and then 40% the experiences, because the thing with migraines is sometimes it is subjective...the fact that you hear it in, like, a story, as well, if it's like a personal narrative, it's a bit better than...a chart saying...a percentage point did this...* [P8]

A total of 10 participants spoke about the need to experience migraine to truly understand it. Being able to read the lived experiences of other people with migraine was beneficial in providing personally relevant information.

> *I mean, my partner's very supportive. But again, there's only so far he can empathize...when I am poorly, because he doesn't know what it's like, he hasn't experienced one before.* [P3]

Catharsis

A total of 10 participants described how they used social media as an outlet for discussing frustrating migraine experiences. Social media was a resource for some participants to cope with the emotions that built up from their experiences:

> *I use it as a tool to cope as it's quite useful for venting.* [P10]

In total, 8 participants described how venting to other people on social media can prevent over-burdening family and friends:

> *If you have a chronic condition...you don't want to bore people absolutely rigid. So it's really nice having that set of strangers who you can actually offload to away from your ordinary life...* [P15]

Another participant described sharing her migraine experience online as *therapeutic*, although 3 participants spoke about how using humor provided some light relief:

> *...one thing that it [Migraine Action Facebook page] had on there was some jokes about migraine...that was quite good, to actually share just silliness about it because sometimes you just have to laugh about it.* [P17]

Theme 4: Presentation and Perception of the Self

Self-Awareness and Awareness of Others

A total of 5 participants described how their identity had evolved since beginning to use social media. This included changes in the way they presented themselves online, as well as changing views of social media uses:

> *[Migraines] weren't part of my identity online because I didn't want to admit that they were a thing...I didn't want them to exist in my life, even though they did and they existed a lot. So that shift in mind-set...bringing them into the conversation...was quite important to me that I was just a bit more attentive on my own identity...* [P13]

In total, 8 participants described how they had different Web-based identities, and some had multiple profiles, often to avoid disclosure of their migraine experience in certain domains of their lives. For some, social media groups had influenced their Web-based behavior and they displayed a different *self* in a private or closed migraine forum, compared with on their personal Facebook timeline:

> *...I haven't posted much at all on my own timeline, it's more on the groups. So I have a different persona on the groups [closed migraine-specific Facebook groups]. "I'm in pain, please help. Having a rubbish day." All that sort of stuff goes on that group, or the groups, and then I tend to put a braver face on to what's going on in everyday life.* [P18]

A total of 5 participants were concerned about how others perceived them on social media, and for some, this affected how they presented themselves on Web-based media:

> *I will put things on there [the closed group] because I wouldn't want my friends and my family to think, "Oh my God, she's going on about migraine again." So I like the fact that it's not public to everybody else.* [P15]

Online and Offline Identity

Several participants alluded to the temporal nature of migraines and how the extremity of acute symptoms could greatly impact their sense of identity and self:

> *...when I'm like well...I'm just somebody who suffers with migraines, whereas when it's extreme, I just feel like I am a migraine.* [P16]

There was also variation in how participants felt social media reflected their migraine experiences. One participant said using Twitter had helped them to continue their role as a knowledgeable scientist after having to leave their job owing to their migraines:

> *it's helped me retain my identity, I guess, after I got sick.* [P4]

Another participant described how social media has helped her to normalize and accept her migraine experience:

> *...the tips I've got have helped me reduce my migraine, to hopefully shove migraine more out of my identity. But now, I don't know whether it is almost desensitization because it comes up on my phone, on the newsfeed, at any time. So that makes it more part of normal life rather than being associated with an acute attack. Maybe in a way it helped it become less of my identity...* [P15]

XSL•FO
RenderX

Discussion

Principal Findings

Our findings suggest that people with migraine are using social media to obtain health-related information to better understand their condition and treatment options. Social media can offer instant access to continuous migraine-related information, as well as social support from empathic others. The opportunity to pool the subjective lived experience of migraines on social media was described as invaluable, and the exchange of support and information was viewed as mutually beneficial.

Some participants viewed social media as an outlet to vent frustrations and validate migraine experiences with other users. They spoke of the invisible and episodic nature of the condition that may contribute to societal misunderstanding about the impact and severity of migraine. Some participants masked their migraine-related behavior using different Web-based sites or closed social media groups to control who saw their migraine-related content. Utilizing social media in this way had enabled participants to retain desired aspects of their identity, depending on how the individual chose to engage with the platform.

Limitations

This study was conducted with a user group of volunteers in 1 country. People who belong to user groups are likely to be more highly educated, just as our volunteers were highly educated people, and are not representative of the whole population. The study also relied on a self-report migraine diagnosis. Future research could sample a broader demographic and use a clinical migraine diagnosis.

Comparison With Previous Work

Many participants said that the information they had accessed on social media enabled them to be more active in their own health care, altering their migraine management and increasing their confidence when interacting with health care professionals. For example, some participants requested treatments that they had learnt about on social media. Increased feelings of control and self-management following social media use have been observed elsewhere in other chronic conditions such as diabetes and hypertension [26,32].

Participants had varied affordances for social media sites. There was a continuum, with more engaged users creating and sharing content (eg, blogging and creating YouTube videos) and those who tended to observe rather than contribute themselves. There was also a group of mid-range users who engage with migraine-specific groups. We have previously observed similar Web-based affordances among people with epilepsy [33].

Benefits of social media use are mostly attributable to social support via Facebook and blogs [34]. Blogging was much less commonly observed in our sample, but Facebook was the most used site. The closed group function on Facebook seemed to create a setting by which people with migraine could overcome invalidation or misunderstanding by identifying with others in the group with whom they had shared experiences. Sharing ideas and information and accessing support within groups were the most common themes in our study, suggesting users benefited from the membership afforded particularly by private social media groups.

Web-Based Migraine Information and Support Group

For some, use of social media for health care purposes may represent an unfulfilled need with their existing care and support networks [35]. The migraine experience is often misunderstood, which can cause isolation and loneliness. The temporality of the condition may restrict an individual's ability to carry out daily activities and participate in their social life. Thus, social media groups are beneficial to those with unpredictable or episodic health conditions, as the support offered is not in a fixed temporal or spatial context. These are key benefits of Web-based support networks that are often restricted in *real life* support group settings [36-38].

An episodic and invisible condition can be difficult to understand by those who do not experience it [3]. Gaining support and information from those with a similar lived experience appears to be a typical use of social media. Overcoming a sense of being alone and sharing experiences have been described as a benefit of attending a self-management group course for people with epilepsy, another invisible, common episodic condition [39]. In a diabetes study about social support from computer-mediated environments, users were more able to engage with their peers and access individualized support tailored to their specific situation [36]. Our findings in migraine are consistent with this.

There is also scope for medical professionals to utilize social media as a health resource, for example, integrating patient data such as electronic diaries in routine practice [40]. Researchers suggest that platforms such as Twitter have the potential to be used for gathering service user feedback [41]. Both user groups and health care providers could learn from this feedback to improve the services and information they provide. Further exploration on the role of health care service providers' use of social media is warranted [41].

Validation of Migraine Experience

Participants spoke about how using social media helped acknowledge their experience of migraine. Migraine is often misclassified or compared with commonplace headache, making migraine seem even more invisible [7]. Other conditions such as chronic fatigue syndrome can also be incorrectly compared with the prevalent state of tiredness. Chronic fatigue patients and fatigued employees reported experiencing more negative interactions with co-workers and insufficient social support compared with people in cancer remission and healthy co-workers [42]. This is perhaps due to lay misconceptions about the condition, lack of clear external illness markers, and parallels drawn with other common conditions. These ideas of similarity challenge the individual with an invisible illness to explain their impairments to feel understood [3]. Thus, the experience of having migraine acknowledged and validated by other users on social media is particularly valuable.

The prevailing opinion that headache is normal and can be alleviated by a tablet may be a barrier to individuals with migraine obtaining adequate support and understanding [32].

XSL•FO
RenderX

Participants in our study often referred to their offline contacts giving them patronizing or unhelpful advice about treating migraines. The sudden onset of migraine along with unpredictable frequency and duration affects people's social, domestic, and professional lives [9,10,43,44]. Some people report feeling that their partner [9] and those in their social network [10] do not believe the impact or severity of their headaches. Associating with a group of others online with similar experiences was described positively by participants in this study, resulting in empowerment and changes in self-management. Social media offered an easily accessible source of understanding from others with a similar lived experience, which offered the added benefit of not overburdening offline support networks.

Identity and Self-Presentation in Migraine

Social context impacts the self that is presented at a family gathering, job interview, or romantic date [5], which is also true of Web-based self-presentation [45]. Goffman compared self-presentation to a stage performance, where the individual chooses what to present to the audience onstage and what to keep private backstage [5]. Self-presentation is likened to a personal exhibit, where individuals act as curators by choosing what to display to their audience [46]. Participants gave varied responses about how migraine fits with their sense of self and self-presentation. Several chose different forms of self-presentation depending on their role on social media, such as being an educator by sharing information with others.

Twitter users have cited their ideal audience on social media to be a mirror image of themselves [45]. Festinger suggested that individuals find their own opinions and personal beliefs more valid when they are shared to a sufficient degree of similarity with a group of others [47]. This desire for similarity may also explain the desire to connect with groups of other users with migraine symptoms. Such connections appeared to empower those who sought them. Our findings indicate that participants engaged with social media as a means of identifying with other users with migraine as a means of talking about issues that they may not have been able to discuss with people outside of these groups. This could be explained using the social identity theory [26], whereby membership to Web-based groups such as this are associated with positive feelings and behaviors [48] within these Web-based communities.

Through such groups, users can share aspects of themselves that they may not do so freely with others in their typical social circles. In part, stigma or a lack of understanding from others may further influence this user behavior in Web-based settings. Belonging to Web-based communities that offer membership based on shared characteristics may help users to form a coherent sense of self that is the same across online and offline encounters [49]. In this way, the access to Web-based groups with positive health-related characteristics may be especially useful to people with conditions requiring support and self-management.

Conclusions

In this study, we broadly enquired how participants with migraine used social media platforms and how this group presented their identity online. Participants described cognitive and affective benefits owing to the use of social media, particularly, gathering and sharing information, as well as receiving ongoing support and validation. They described learning about new therapies and side effects that had helped them in consultations with doctors. Participants particularly valued the affordances provided by Facebook groups.

Similar to others with potentially stigmatizing conditions, people with migraine want empathy and understanding for their subjective, but real, disability [6,39]. Based on our findings, participants described protecting themselves, or this aspect of their identity, from people who might disparage or reject them because of their condition. The apparent secrecy offered by social media helped them come to terms and cope with their condition, while simultaneously and separately managing the rest of their lives.

Social media can help validate the experience of migraine and in turn help people who experience migraines to feel better understood and less alone. How migraine is part of a person's identity and represented online varies. Further understanding about the needs of people with long-term chronic conditions may help in the development of future Web-based interventions to improve health and well-being. Activity that occurs in closed groups helps people to accept and manage their condition in a separate but synergistic way.

Official Institutions for further information on migraine are as follows: International Headache Society [50], European Headache Federation [51], American Migraine Foundation [52], American Pain Society [53], European Headache Alliance [54], National Headache Foundation [55], Migraine Trust [56], Migraine Action [57].

Acknowledgments

The authors would like to particularly thank Gabriella Wojewodka and Raphael Underwood, the Migraine Trust and Migraine Action for advertising the project, and the study participants. This paper has been made possible as a result of funding from the European Research Council for the project, Ego-Media: The Impact of New Media on Forms and Practices of Self-Presentation (FP7/2007-2013; grant agreement #340331).

Authors' Contributions

LR designed the study and supervised the team throughout. RS, PK, and CP contributed to data collection. All authors contributed to data analysis and interpretation. CP, LR, and AM wrote the final manuscript, with input from other authors.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Weatherall MW. The diagnosis and treatment of chronic migraine. Ther Adv Chronic Dis 2015 May;6(3):115-123 [FREE Full text] [doi: 10.1177/2040622315579627] [Medline: 25954496]
2. Goadsby PJ, Sprenger T. Current practice and future directions in the prevention and acute management of migraine. Lancet Neurol 2010 Mar;9(3):285-298. [doi: 10.1016/S1474-4422(10)70005-3] [Medline: 20170842]
3. Lingsom S. Invisible impairments: dilemmas of concealment and disclosure. Scand J Disability Res 2008 Mar;10(1):2-16. [doi: 10.1080/15017410701391567]
4. Goffman E. Stigma: Notes On The Management Of Spoiled Identity. London, England: Penguin Books, Limited; 1990.
5. Goffman E. The Presentation Of Self In Everyday Life. London, England: Penguin Books, Limited; 1990.
6. Steiner T, Stovner L, Vos T, Jensen R, Katsarava Z. Migraine is first cause of disability in under 50s: will health politicians now take notice? J Headache Pain 2018 Dec 21;19(1):17 [FREE Full text] [doi: 10.1186/s10194-018-0846-2] [Medline: 29468450]
7. Lonardi C. The passing dilemma in socially invisible diseases: narratives on chronic headache. Soc Sci Med 2007 Oct;65(8):1619-1629. [doi: 10.1016/j.socscimed.2007.07.007] [Medline: 17716794]
8. Kleinman A. The Illness Narative: Suffering, Health and the Human Condition. Cambridge: Basic Books; 1988.
9. Buse D, Scher A, Dodick DW, Reed ML, Fanning KM, Manack Adams A, et al. Impact of migraine on the family: perspectives of people with migraine and their spouse/domestic partner in the CaMEO study. Mayo Clin Proc 2016 Apr 25;91:596-611 [FREE Full text] [doi: 10.1016/j.mayocp.2016.02.013] [Medline: 27132088]
10. Malone C, Bhowmick A, Wachholtz A. Migraine: treatments, comorbidities, and quality of life, in the USA. J Pain Res 2015;8:537-547 [FREE Full text] [doi: 10.2147/JPR.S88207] [Medline: 26316804]
11. Bigal M, Serrano D, Reed M, Lipton R. Chronic migraine in the population: burden, diagnosis, and satisfaction with treatment. Neurology 2008 Aug 19;71(8):559-566. [doi: 10.1212/01.wnl.0000323925.29520.e7] [Medline: 18711108]
12. Peters M, Abu-Saad H, Vydelingum V, Dowson A, Murphy M. Migraine and chronic daily headache management: a qualitative study of patients' perceptions. Scand J Caring Sci 2004 Sep;18(3):294-303. [doi: 10.1111/j.1471-6712.2004.00279.x] [Medline: 15355524]
13. Nascimento T, DosSantos M, Danciu T, DeBoer M, van Holsbeeck H, Lucas S, UMSoD (Under)Graduate Class of 2014, et al. Real-time sharing and expression of migraine headache suffering on Twitter: a cross-sectional infodemiology study. J Med Internet Res 2014 Apr 3;16(4):e96 [FREE Full text] [doi: 10.2196/jmir.3265] [Medline: 24698747]
14. Kaplan A, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. Bus Horiz 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
15. Boyd D, Ellison N. Social network sites: definition, history, and scholarship. J Comput-mediat Comm 2007;13(1):210-230. [doi: 10.1111/j.1083-6101.2007.00393.x]
16. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368 [FREE Full text]
17. Moorhead S, Hazlett D, Harrison L, Carroll J, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
18. Tajfel H, Turner J. The social identity theory of intergroup behavior. In: Jost J, Sidanius J, editors. Political Psychology: Key Readings. New York: Psychology Press; 2004:276-293.
19. Hogg M, Reid S. Social identity, self-categorization, and the communication of group norms. Commun Theor 2006;16(1):7-30. [doi: 10.1111/j.1468-2885.2006.00003.x]
20. Tafjel H, Turner TJ. An integrative theory of intergroup conflict. In: Austin WG, Worchel S, editors. The Social Psychology of Intergroup Relations. Monterey, Calif: Brooks Cole Publishing; 1979:33-47.
21. Tajfel H. Social identity and intergroup behaviour. Soc Sci Inf 1974 Apr 1;13(2):65-93. [doi: 10.1177/053901847401300204]
22. Brivio E, Ibarra FC. Self presentation in blogs and social networks. Stud Health Technol Inform 2009;144:113-115. [Medline: 19592744]
23. Mosen D, Schmittdiel J, Hibbard J, Sobel D, Remmers C, Bellows J. Is patient activation associated with outcomes of care for adults with chronic conditions? J Ambul Care Manage 2007;30(1):21-29. [Medline: 17170635]
24. Lorig K, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
25. Merolli M, Gray K, Martin-Sanchez F. Health outcomes and related effects of using social media in chronic disease management: a literature review and analysis of affordances. J Biomed Inform 2013 Dec;46(6):957-969 [FREE Full text] [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]

XSL•FO
RenderX

26. Grosberg D, Grinvald H, Reuveni H, Magnezi R. Frequent surfing on social health networks is associated with increased knowledge and patient health activation. J Med Internet Res 2016 Dec 10;18(8):e212 [FREE Full text] [doi: 10.2196/jmir.5832] [Medline: 27511272]
27. Linnman C, Maleki N, Becerra L, Borsook D. Migraine tweets - what can online behavior tell us about disease? Cephalalgia 2013 Jan;33(1):68-69 [FREE Full text] [doi: 10.1177/0333102412465207] [Medline: 23136208]
28. Morgan M, Jenkins L, Ridsdale L. Patient pressure for referral for headache: a qualitative study of GPs' referral behaviour. Br J Gen Pract 2007 Jan;57(534):29-35 [FREE Full text] [Medline: 17244421]
29. Morgan M, Cousins S, Middleton L, Warriner-Gallyer G, Ridsdale L. Patients' experiences of a behavioural intervention for migraine headache: a qualitative study. J Headache Pain 2015;17:16 [FREE Full text] [doi: 10.1186/s10194-016-0601-5] [Medline: 26922332]
30. Ritchie JS, Spencer L, O'Connor W. Carrying out qualitative analysis. In: Ritchie JS, Lewis J, editors. Qualitative Research Practice: A Guide For Social Science Students And Researchers. London: Sage Publications Ltd; 2003:219-262.
31. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
32. Allen C, Vassilev I, Kennedy A, Rogers A. Long-term condition self-management support in online communities: a meta-synthesis of qualitative papers. J Med Internet Res 2016 Mar 10;18(3):e61 [FREE Full text] [doi: 10.2196/jmir.5260] [Medline: 26965990]
33. McKinlay A, Ridsdale L. Views of people with epilepsy about web-based self-presentation: a qualitative study. Interact J Med Res 2018 Dec 21;7(2):e10349 [FREE Full text] [doi: 10.2196/10349] [Medline: 30578229]
34. Patel R, Chang T, Greysen S, Chopra V. Social media use in chronic disease: a systematic review and novel taxonomy. Am J Med 2015 Dec;128(12):1335-1350. [doi: 10.1016/j.amjmed.2015.06.015] [Medline: 26159633]
35. Smailhodzic E, Hooijsma W, Boonstra A, Langley D. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Dec 26;16:442 [FREE Full text] [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
36. Lewinski A, Fisher E. Social interaction in type 2 diabetes computer-mediated environments: how inherent features of the channels influence peer-to-peer interaction. Chronic Illn 2016 Jun;12(2):116-144. [doi: 10.1177/1742395315601414] [Medline: 26304692]
37. Donath J. Signals in Social Supernets. J Comput Mediat Comm 2007;13(1):231-251. [doi: 10.1111/j.1083-6101.2007.00394.x]
38. Linnman C, Maleki N, Becerra L, Borsook D. Migraine tweets - what can online behavior tell us about disease? Cephalalgia 2013 Jan;33(1):68-69 [FREE Full text] [doi: 10.1177/0333102412465207] [Medline: 23136208]
39. Ridsdale L, Philpott S, Krooupa A, Morgan M. People with epilepsy obtain added value from education in groups: results of a qualitative study. Eur J Neurol 2017 Dec;24(4):609-616 [FREE Full text] [doi: 10.1111/ene.13253] [Medline: 28181344]
40. Allena M, Cuzzoni M, Tassorelli C, Nappi G, Antonaci F. An electronic diary on a palm device for headache monitoring: a preliminary experience. J Headache Pain 2012 Oct;13(7):537-541 [FREE Full text] [doi: 10.1007/s10194-012-0473-2] [Medline: 22842873]
41. Shepherd A, Sanders C, Doyle M, Shaw J. Using social media for support and feedback by mental health service users: thematic analysis of a twitter conversation. BMC Psychiatry 2015 Feb 19;15:29 [FREE Full text] [doi: 10.1186/s12888-015-0408-y] [Medline: 25881089]
42. Prins J, Bos E, Huibers MJ, Servaes P, van der Werf SP, van der Meer JW, et al. Social support and the persistence of complaints in chronic fatigue syndrome. Psychother Psychosom 2004;73(3):174-182. [doi: 10.1159/000076455] [Medline: 15031590]
43. Young W, Park J, Tian I, Kempner J. The stigma of migraine. PLoS One 2013;8(1):e54074 [FREE Full text] [doi: 10.1371/journal.pone.0054074] [Medline: 23342079]
44. Cooke L, Becker W. Migraine prevalence, treatment and impact: the Canadian women and migraine study. Can J Neurol Sci 2010 Sep;37(5):580-587. [Medline: 21059502]
45. Marwick A, boyd D. I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. New Media Soc 2010 Jul 7;13(1):114-133. [doi: 10.1177/1461444810365313]
46. Hogan B. The presentation of self in the age of social media: distinguishing performances and exhibitions online. Bull Sci Technol Soc 2010 Nov;30(6):377-386. [doi: 10.1177/0270467610385893]
47. Festinger L. Informal social communication. Psychol Rev 1950 Sep;57(5):271-282. [Medline: 14776174]
48. Hogg M, Terry D, White K. A tale of two theories: a critical comparison of identity theory with social identity theory. Soc Psychol Q 1995 Dec;58(4):255 [FREE Full text] [doi: 10.2307/2787127]
49. Manago A, Graham M, Greenfield P, Salimkhan G. Self-presentation and gender on MySpace. J Appl Dev Psychol 2008 Nov;29(6):446-458. [doi: 10.1016/j.appdev.2008.07.001]
50. International Headache Society. URL: http://www.ihs-headache.org/ [accessed 2018-06-15]
51. European Headache Federation. URL: https://ehf-org.org/ [accessed 2018-06-15]
52. American Migraine Foundation. URL: https://americanmigrainefoundation.org/ [accessed 2018-06-15]
53. American Pain Society. URL: http://americanpainsociety.org/ [accessed 2018-06-15]
54. European Headache Alliance. URL: https://www.emhalliance.org/ [accessed 2018-06-15]

XSL•FO
RenderX

55. National Headache Foundation. URL: https://headaches.org/ [accessed 2018-06-15]
56. Migraine Trust. URL: https://www.migrainetrust.org/ [accessed 2018-06-15]
57. Migraine Action. URL: http://www.migraine.org.uk/information/ [accessed 2018-06-15]

Abbreviations

B: blogs
F: Facebook
H: Health Unlocked
I: Instagram
M: Migraine.com
MT: My Migraine Team
NHS: National Health Service
P: Pinterest
T: Twitter
V: Vestibular migraine website with online community
Y: YouTube

Edited by G Eysenbach; submitted 22.03.18; peer-reviewed by P Martelletti, Y Zhang; comments to author 26.04.18; revised version received 15.06.18; accepted 18.04.19; published 19.06.19.

Please cite as:
Pearson C, Swindale R, Keighley P, McKinlay AR, Ridsdale L
Not Just a Headache: Qualitative Study About Web-Based Self-Presentation and Social Media Use by People With Migraine
J Med Internet Res 2019;21(6):e10479
URL: http://www.jmir.org/2019/6/e10479/
doi:10.2196/10479
PMID:31219049

©Carly Pearson, Rosanna Swindale, Peter Keighley, Alison Ruth McKinlay, Leone Ridsdale. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mechanisms of Social Media Effects on Attitudes Toward E-Cigarette Use: Motivations, Mediators, and Moderators in a National Survey of Adolescents

Hyunyi Cho[1], PhD; Wenbo Li[1], MA; Lijiang Shen[2], PhD; Julie Cannon[3], MA

[1]School of Communication, The Ohio State University, Columbus, OH, United States

[2]Department of Communication Arts and Sciences, Pennsylvania State University, University Park, PA, United States

[3]Department of Communication, Cornell University, Ithaca, NY, United States

Corresponding Author:
Hyunyi Cho, PhD
School of Communication
The Ohio State University
154 N Oval Mall
Columbus, OH,
United States
Phone: 1 614 292 3400
Email: cho.919@osu.edu

Abstract

Background: Exposure to risk behavior on social media is associated with risk behavior tendencies among adolescents, but research on the mechanisms underlying the effects of social media exposure is sparse.

Objective: This study aimed to investigate the motivations of social media use and the mediating and moderating mechanisms of their effects on attitude toward electronic cigarette (e-cigarette) use among adolescents.

Methods: Using data from a national sample survey of adolescents (age=14-17 years, N=594), we developed and validated a social media use motivation scale. We examined the roles of motivations in the effect of social media use on risk exposure and risk attitude.

Results: Motivations for social media use included agency, self-expression, realism, social learning, social comparison, and filter. These motivations were associated differentially with the frequency of use of Facebook, Instagram, Snapchat, and YouTube. Frequency of social media use was positively associated with exposure to e-cigarette messages across the four platforms (*P*s<.001). Exposure to e-cigarette messages on Instagram (*P*=.005) and Snapchat (*P*=.03) was positively associated with attitude toward e-cigarette use. Perceived social media realism moderated the effects of e-cigarette message exposure such that when realism was high, the exposure effect was amplified, but when realism was low, the effect was mitigated (*P*<.001). A three-way interaction effect (*P*=.02) among exposure, social learning motivation, and social norm on attitude toward e-cigarette use was found. When perceived social norm was high, the moderating effect of social learning motivation on e-cigarette use attitude was amplified, but when social norm was low, the social learning motivation effect was attenuated.

Conclusions: Because perceived social media realism moderates the effect of exposure to e-cigarette messages on attitude toward e-cigarette use, future intervention efforts should address the realism perceptions. The three-way interaction among exposure, social learning motivation, and social norm indicates the importance of addressing both the online and offline social environments of adolescents. The social media use motivation scale, reflecting perceived affordances, is broadly applicable. Understanding social media use motivations is important, as they indirectly influence attitude toward e-cigarette use via frequency of social media use and/or frequency of exposure to e-cigarette messages on social media.

(J Med Internet Res 2019;21(6):e14303) doi:10.2196/14303

KEYWORDS

adolescents; e-cigarettes; motivation; affordances; agency; realism; self-expression; social comparison; social learning; social media; filter; uses and gratifications

XSL•FO
RenderX

Introduction

Background

Exposure to risk-related content on social media has been associated with positive attitudes toward risk behavior and adoption of the risk behavior, including alcohol use [1] and drug use [2]. Research shows that adolescents are particularly vulnerable to such social media influences, as the generation is growing up with social media as an important source and channel of socialization and as a social environment [3]. According to a Pew Research Center survey [4], the most popular platforms among adolescents were YouTube, followed by Instagram, Snapchat, and Facebook.

A particularly concerning risk behavior among adolescents is e-cigarette use. Although recent research shows that e-cigarette use is more efficacious than nicotine-replacement products in helping cigarette smokers quit [5], uptake of e-cigarette use can be harmful for adolescents. Research has found that e-cigarettes contain nicotine [6], which adversely affects brain development up to the age of 25 years [7]. Nicotine use during adolescence elevates the risk of addiction to other drugs [8]. Furthermore, the aerosol from e-cigarettes exposes both users and bystanders to harmful substances [9]. For these reasons, in 2018, the US Surgeon General called for urgent and aggressive action to protect young people from the harm, declaring youth e-cigarette use an epidemic [10].

Although findings connecting social media exposure and risk behavior including e-cigarette use have been increasing, sparse research has investigated the mechanisms with which the effects occur. The goal of this study was to address this gap in knowledge by investigating *how* social media influences adolescent e-cigarette use. When the processes underlying the connection between exposure and effect are identified, strategies for effective interventions could be developed. Therefore, this study sought to identify motivations of social media use and to investigate their mediating and moderating roles in the association between social media–based risk exposure and risk behavior attitude among adolescents.

Conceptual Bases

In this study, we conceptualize social media use as a motivated behavior driven by individuals' desire to fulfill their psychological and social needs [11] and hypothesize that differential motivations influence individuals' choice of differential social media platforms, as they offer differential affordances. Frequency of use of differential platforms will be differentially associated with e-cigarette message exposure, which, in turn, influences attitude toward e-cigarette use.

To understand the effects of social media use on risk behavior attitudes, the motivations for social media use should be examined first. The uses and gratifications framework [11] explains that people use media to fulfill their individual needs and that there are motivations associated with people's choice to consume media. For a given medium use (eg, television), there can be a range of motivations. Television use motivations, for example, include learning, relaxation, companionship, escape, arousal, and passing time [12,13]. Building on the uses and gratifications paradigm, scholars have advanced the uses and effects perspective, which predicts that differential motivations of media use will lead to differential effects [13,14]. Despite the theoretical advance, little research has empirically examined the relationship between motivations and effects, especially with respect to motivations, uses, and effects of social media.

As the media landscape has changed from mass media including television to the internet to Web 2.0, researchers have investigated the motivations associated with the use of various social media platforms. Studies have examined motivations or gratifications associated with the use of Facebook [15], Twitter [16], Instagram [17], YouTube [18], and Snapchat [19]. Although the information about motivations of using each social media platform is valuable, as they may be variable across the platforms, the motivations may also share commonalities. It may be more advantageous to study new media as a mix of attributes rather than discrete entities [20], as identifying the core motivations that can be differentially applied to different platforms may be of greater generative utility than identifying platform-specific motivations.

Moreover, we focus on the unique characteristics of Web 2.0 in this research. The existing conceptualizations and measurement of social media use motivations draw on the typology that has been used for traditional mass media use [21]. Observing that the available descriptions of new media gratifications may be more general than the nuanced gratifications available through new media, Sundar and Limperos [21] suggested possible new gratifications from new media, anchoring them to four classes of features: modality, agency, interactivity, and navigability.

Finally, extant research has yet to take into account the aspects of user-generated content. In mass media, users did not have a chance to contribute content, as the mass media content has been created by professionals. In contrast, Web 2.0, on which various social media platforms are based, relies on user-generated content and collective knowledge and sentiment [22,23]. Three interdependent gratifications available from user-generated media include consumption, participation, and production [22]. This study seeks to capture this new media ecology available on Web 2.0.

Moreno and colleagues' [24] Facebook influence model captures a number of new gratifications linked to Facebook use, some of which could be extended to other social media use. Through a conceptual mapping approach, they identified 10 facets of Facebook use gratifications, including connection to people, far reaching, fast communication, curiosity about others, business and promotion, accessibility/adaptability, data/information, social norm establishment, identity expression, influence on identity, distraction, positive experience, and negative experience [24].

Motivations for Social Media Use

Based on the past research reviewed above, we aimed to identify core motivations that may underlie the use of divergent social media platforms and may explain the social and psychological processes of social media effects. These motivations included

XSL•FO
RenderX

agency, self-expression, realism, social learning, social comparison, and filter. The attributes and features available on various social media platforms may differentially accommodate these motivations. Notably, as discussed below, the social aspects of the motivations are frequently networked in nature, and the psychological motivations contribute to collective intelligence and sentiment on Web 2.0.

The agency and self-expression motivations represent the identity-integral function of social media. Not included in motivations ascertained for the use of mass media, the abovementioned motivations indicate a unique aspect of Web 2.0, which is often characterized as participative and user-generated. Agency refers to the motivation to influence others by sharing one's own ideas and messages. Relevant to the agency motivation, research found that the expressive and performative involvement in user-generated content on the internet facilitated online and offline political participation among adolescents [25,26]. Agency was included as a gratification integral to new media uses [21] and Facebook use [24].

The self-expression motivation reflects a gratification sought and obtainable from user-generated media, unavailable from mass media. On user-generated media, the desire for production intersects with the desire for participation and consumption [22]. The internet makes it possible for individuals to experiment with their identities [27], and youth express their identities on social media while concurrently being open to social media's influences on these identities [24]. Creativity was one of the motivations associated with young adults' use of Instagram [17].

Perceived media realism has been an important explanatory variable of media effects [28]. In this study, however, we conceptualize social media realism differently from previous research on mass media realism. With traditional media, the industry and professionals determined the content. In social media, users are creators of the content as well as its consumers. Therefore, we developed items to capture this unique aspect of social media content and users' assessment of the content created by other users.

The media provide a powerful apparatus to learn about the world and others. Social learning has been an important motivation for using media [12,13]. Mass media research found, for example, that soap opera viewers with information motivation reported greater attention to the program and engagement with characters during viewing than viewers with the motivation to merely pass time [29]. Social learning motivation has been linked to the use of Facebook [30] and Instagram [17]. Of note, in this study, we assert that the "social" in social learning motivation for using social media differs from the motivation for using mass media. In social media, the "social" is frequently networked in nature. On the other hand, in mass media, the learning motivation is more societal than social, as the target of learning is more diffusive in nature.

With metrics such as likes, ratings, and positive and negative comments that were unavailable in mass media, social media provides insight on others' behaviors and attitudes. With these features, social media may provide knowledge about social norms [21,24,31] and a benchmark for social comparison [32]. Adolescents' social comparison activities in social media affect their identity development under certain conditions [33].

In addition, social media allows users to construct their own world by using the function of filters [21,34], which may provide psychological contentment and comfort functions but could also obstruct a more comprehensive and objective perspective on what is going on in society [35]. As in this virtual community, only like-mined people may congregate together, reinforcing and strengthening shared viewpoints, concerns have been raised that social media filters can facilitate social fragmentation and polarization [36,37].

With regard to the motivation mechanisms, we propose the following hypothesis:

> *Hypothesis 1: Social media motivations include agency, self-expression, perceived realism, social learning, social comparison, and filter.*

Mediating Mechanisms

The abovementioned motivations may reflect perceived affordances of social media. Affordances are "action possibilities" toward which the stimuli in the environment suggest that humans act [38]. The concept of perceived affordances refers to user experiences and evaluations rather than the features themselves [39]. As each social media platform offers a mix of features, use of these platforms comes with convergent and divergent gratifications. For example, some of the Facebook use gratifications (eg, identity expression) [24] overlap with the Instagram use gratifications (eg, creativity) [17]. Communication channels including phone, email, texting, Facebook, and Snapchat were perceived to provide differential levels of social affordances [40]. Therefore, the mix of motivations is likely associated differentially with the use of differential social media platforms. The frequency of using social media, in turn, may predict the probability of exposure to e-cigarette advertisements and posts. At present, social media is the main outlet of e-cigarette marketing activities [41-43]. Social media–based exposure to e-cigarette messages was positively associated with e-cigarette expectancies among young adults [44]. Extending this prior research, we aim to investigate the association between social media e-cigarette exposure and attitudes toward e-cigarette use in a population of adolescents. Furthermore, we seek to examine the process of the exposure effects.

With regard to the mediating mechanisms, we propose the following hypotheses:

> *Hypothesis 2a: Social media use motivations are differentially associated with the frequency of use of differential social media platforms.*
>
> *Hypothesis 2b: Social media use frequency is positively associated with exposure to e-cigarette messages on social media.*
>
> *Hypothesis 2c: Exposure to e-cigarette messages on social media is positively associated with attitude toward e-cigarette use.*

XSL•FO
RenderX

Moderating Mechanisms

Two-Way Interaction

In addition to indirectly influencing e-cigarette message exposure and e-cigarette use attitude, social media use motivations may moderate the effect of the exposure on attitude, that is, the effects of risk exposure on social media can be attenuated or amplified depending on the motivations of social media use. One of the factors that could modulate harmful social media effects may be realism judgment. The more users believe that other user-generated content is a representation of their true self and a truthful depiction of their beliefs, emotions, and lives, the stronger are the effects of the exposure to the content. On the other hand, if the users think that social media representations deviate from users' true selves and lives, the effects of the exposure would be mitigated.

Three-Way Interaction

The social dimensions of social media use motivation (eg, social learning, social comparison, and filter) may also moderate the effects of risk exposure on social media. Importantly, because of the networked nature of the social world represented in one's social media environment, the influence of the social motivations may be qualified by social norm. For example, while those who use social media with stronger social learning motivation are more likely to be affected by prevalent or glamorous depictions of risk behavior than those with weaker social learning motivation, the influence of social learning motivation may be further moderated by the kind of social world that one maintains. If, for example, friends in one's social network use e-cigarettes, then the effect of social learning motivation may be amplified. However, if no one in one's social network uses e-cigarettes, then the effect of social learning motivation may be mitigated.

Regarding the moderating mechanisms, we proposed the following hypotheses:

> *Hypothesis 3a: The effect of exposure on attitude is moderated by realism motivation such that high realism amplifies the exposure effect on attitude and low realism attenuates it.*
>
> *Hypotheses 3b-d: The effect of exposure on attitude is moderated by social learning (b), social comparison (c), and filter motivations (d), which are, in turn, moderated by social norm. When social norm is high, the moderating effects of these motivations on attitude are amplified, but when social norm is low, the effects of these motivations are attenuated.*

Methods

Design

This study was conducted as part of a larger project investigating adolescent social media use and risk behavior. Participants answered questions about their social media use, risk exposure on social media, and attitude toward e-cigarette use prior to seeing a 1-minute anti–e-cigarette video message. The items assessing motivations associated with social media use were given at the end of the study. To prevent exposure to the 1-minute video message from having any effect on the motivation measures, a battery of items, including those about other adolescent risk behaviors, was assessed after the exposure prior to the motivation assessment.

Participants

Participants (N=594) were adolescents aged 14-17 (mean 15.48, SD 1.12) years recruited through Ipsos (formerly GfK), a survey firm providing a probability-based sample of the US population. Male and female adolescents comprised 47.4% and 52.6% of the sample, respectively. The majority of the adolescents were white individuals (65.0%); the rest were Hispanic (16.4%), black (7.6%), other (6.4%), and mixed race (4.6%) individuals. Both parental consent and adolescent assent were obtained prior to the study.

Measures

Social media use motivations were measured using the scale designed for this study. Items were developed to measure the posited six dimensions of social media use motivations. They were a mixture of items taken from Moreno et al [24] and Sundar and Limperos [21] and those that were developed in this study. For the dimensions of agency, social comparison, and filter, items were taken from the studies of Moreno et al [24] and Sundar and Limperos [21]. Items for the self-expression dimension were taken from Moreno et al [24]. One item of the realism dimension was taken from Sundar and Limperos [21]. Multimedia Appendix 1 presents a full description of the social media use motivation measures. The response scale ranged from 1 - "strongly disagree" to 5 - "strongly agree." Table 1 presents means, SDs, and reliability alphas of and correlations between the dimensions.

Regarding social media use, the frequency of use of the four most popular social media platforms among adolescents [4] was assessed. These platforms included Facebook, Instagram, Snapchat, and YouTube. The response scale ranged from 1 - "I don't use this" to 7 - "five or more times a day." Table 2 presents the descriptive data.

Exposure to e-cigarette messages on social media was measured by asking adolescents how often they saw (1) e-cigarette advertisements and (2) pictures in which people were using e-cigarettes on each of the aforementioned four social media platforms. These two items assessing exposure to ads and usage images were averaged to create an index for each platform. The two items were substantially correlated (r=0.75, 0.66, 0.63, and 0.65 for Facebook, Instagram, Snapchat, and YouTube, respectively). The response scale ranged from 1 - "never" to 5 - "very often." Table 2 shows the distribution.

Attitude toward e-cigarette use was measured using the Osgood semantic differential scale [45]. Specifically, three pairs of bipolar adjectives including bad/good, undesirable/desirable, and unfavorable/favorable were provided and scored on a 5-point scale. Higher scores indicated more positive attitude.

The social norm measure was focused on friends' descriptive norm and measured using the Fishbein scale [46] with the question "how many of your friends use e-cigarettes?" The response scale ranged from 1 - "none" to 4 - "all of them."

XSL•FO
RenderX

Table 1. Means, SDs, reliability (α), and correlations (r) between social media motivations.

#	Motivation	Mean (SD)	α	r					
				1	2	3	4	5	6
1	Agency	3.29 (1.00)	0.94						
2	Filter	3.14 (0.81)	0.79	0.62					
3	Self-expression	3.13 (0.97)	0.91	0.78	0.58				
4	Social learning	3.41 (0.86)	0.91	0.68	0.71	0.61			
5	Social comparison	3.21 (0.94)	0.90	0.61	0.70	0.54	0.75		
6	Realism	2.69 (0.81)	0.84	0.52	0.60	0.54	0.49	0.50	

Table 2. Social media use and exposure to e-cigarette messages on social media among adolescents. All values are given as mean (SD) scores.

Factor	Facebook	Instagram	Snapchat	YouTube
General social media use	2.80 (2.20)	4.42 (2.53)	4.24 (2.63)	5.39 (1.80)
Exposure to e-cigarette messages	1.46 (0.77)	1.81 (0.89)	1.81 (0.93)	1.98 (0.92)

Results

Analysis Strategy

We employed confirmatory factor analyses to test hypothesis 1. Linear regression, PROCESS macro of Hayes [47], and bootstrapping estimation approach of Shrout and Bolger [48] were used to test hypotheses 2 and 3.

Hypothesis 1

Hypothesis 1 predicted that social media motivations would include the dimensions of agency, self-expression, realism, social learning, social comparison, and filter. Confirmatory factor analyses were conducted to test this hypothesis. Unless unidimensionality of a scale is established for the first-order factors, evidence in support of its unidimensionality should be obtained from two sources: (1) a first-order oblique multifactor model that should fit the data and the correlations among the first-order factors should be similar, and (2) statistical equivalence has to be established between the first-order multifactor model and a second-order single-factor model. The degree of freedom for the first-order single-factor model (20 items, 1 factor) was 170, that for the first-order multifactor model (20 items, 6 associated factors) was 164, and that for the second-order single-factor model was 166. With these parameters and assuming α=0.05, a sample size of 568 yielded statistical power of >0.99 when testing these factor models [49].

Input and model specifications are as follows. Individuals' responses to the 20 social media use items were submitted to LISREL 8.80 (Scientific Software International, Inc, IL) for confirmatory factor analyses. First, a first-order single-factor model was estimated, where all 20 items were specified to load on one latent factor. Second, a first-order oblique six-factor model was estimated, where the 6 first factors were allowed to be associated with each other. Third, a second-order single-factor model was estimated, where the second-order factor loaded on the 6 first-order factors, which were not allowed to be correlated.

To evaluate the overall fit of the models to the data, four fit indices were considered. First, the goodness-of-fit index (GFI) produces values ranging from 0 to 1, with values in excess of 0.90 indicating good fit. Second, the comparative fit index (CFI) produces values ranging from 0 to 1, with values larger than 0.90 indicating good fit. Third, Browne and Cudeck [50] contend that values of the root mean square error of approximation (RMSEA) of 0.08 or lower indicate reasonable fit, although values of 0.06 or below should be preferred. Fourth, the Bayesian information criterion (BIC) is constructed such that negative values provide evidence of model fit, while positive BIC values suggest problematic model fit. Differences in BIC of 2 are thought to provide some evidence; ≥6, strong evidence; and ≥10, very strong evidence for the superiority of the model with a more negative BIC value over another model [51].

The first-order single-factor model did not fit the data; the values of this model are as follows: χ^2_{170}= 3147.97, *P*<.001, RMSEA=0.21, GFI=0.57, CFI=0.86, and BIC=2061.06. The first-order oblique six-factor model was a good fit to the data: χ^2_{155}=639.18, *P*<.001, RMSEA=0.08, GFI=0.90, CFI=0.98, and BIC=–351.83. Additional evidence was obtained from (1) the standardized factor loadings (the three factors had similar and reasonably high loadings on the indicators, ranging from 0.76 to 0.92 and similar within each factor) and (2) the substantive correlations among the six factors (ranging from 0.54 to 0.84), thereby providing clear indication of nonorthogonality.

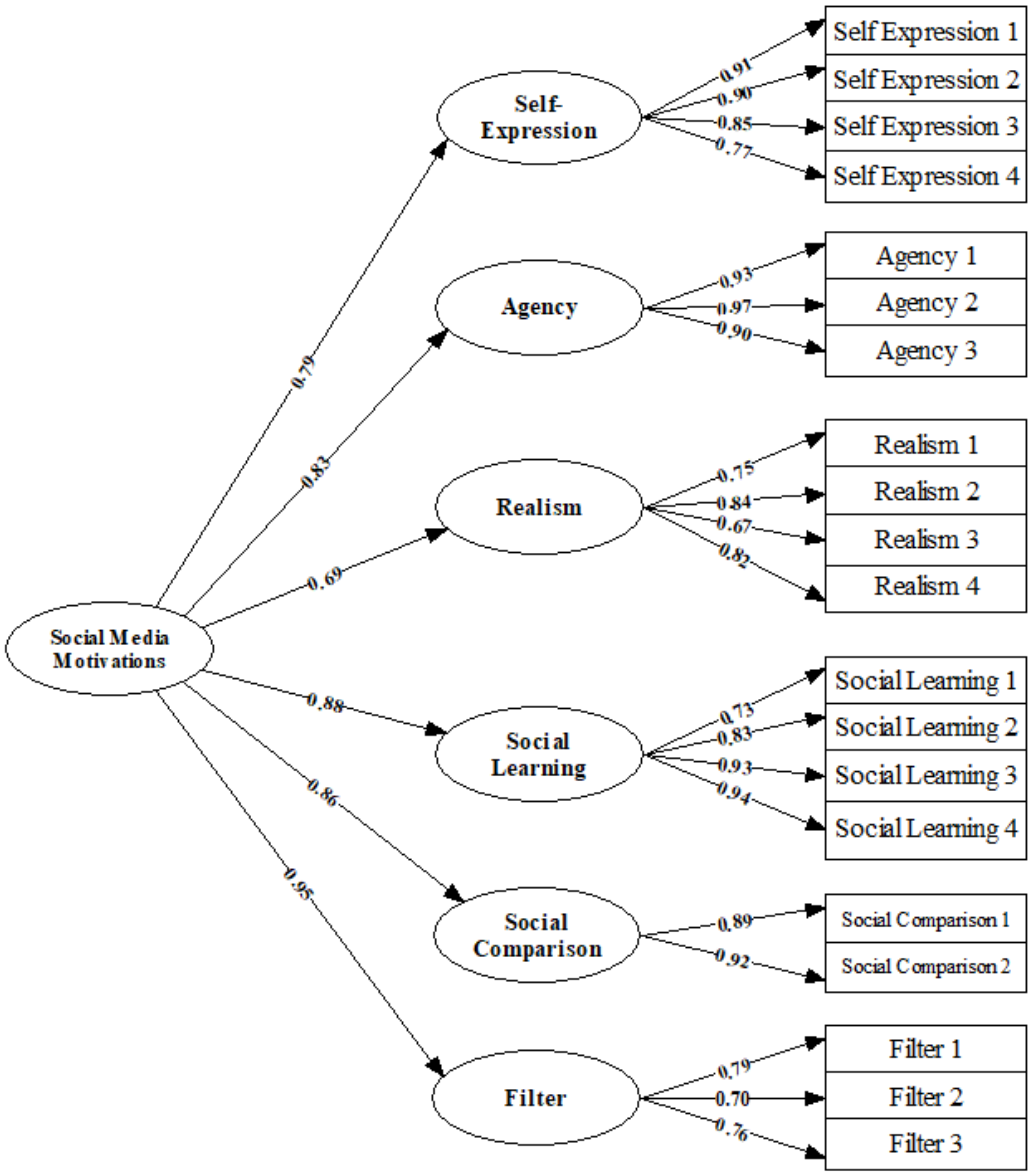
The second-order single-factor model was nested within the oblique first-order six-factor model. It was not surprising that the second-order single-factor model yielded worse fit than the first-order six-factor model: χ^2_9= 211.69 and *P*<.001. However, given the exceptional statistical power (ie, >0.99) in model testing, we considered this discrepancy to be acceptable. The absolute indices showed that the second-order single-factor model was also a good fit to the data: χ^2_{164}=850.87, *P*<.001, RMSEA=0.08, GFI=0.90, and CFI=0.98, except for BIC=–197.68. The values for RMSEA, GFI, and CFI were

XSL•FO
RenderX

almost identical between the second-order single-factor model and the first-order six-factor model. Together, these values indicated that the second-order single-factor model provided a plausible account of the data. However, the Chi-square test and BIC difference of 156.15 in favor of the first-order six-factor oblique model suggested potential variances in the factor structure of social media motivations.

The factor loadings of the six first-order factors on the second order factor provided additional support: The factor loadings were 0.78 for self-expression, 0.82 for agency, 0.68 for realism, 0.87 for social learning, 0.85 for social comparison, and 0.94 for filter. These results provided evidence that the second-order single-factor model was adequate for the social media use motivation scale and could be considered equivalent to the first-order six-factor model. Figure 1 presents the factor structure.

Figure 1. Factor structure.



Hypothesis 2

Hypothesis 2a predicted that social media use motivations would be associated differentially with different social media platform usage. For each of the four social media platforms, the frequency of use was regressed onto the six social media use motivations and the covariates of adolescent age, sex, and race. As Table 3 summarizes, the motivations were associated differentially across the four platforms. Facebook use was associated with the motivations of agency (P=.005) and filter (P=.008). Instagram use was associated with self-expression (P<.001) and social learning motivations (P=.009). Likewise, Snapchat use was associated with self-expression (P<.001) and social learning (P=.01). YouTube use was associated with social comparison motivation (P=.002). In this regression model, where the motivations were controlling for each other, realism was unrelated to any of the social media platform use. However, without controlling for other motivations (but controlling for age, sex, and race), realism was associated with the use of Facebook (β=0.209), Instagram (β=0.201), and Snapchat (β=0.177; all P<.001), but not YouTube (β=0.071, P=.09). Similarly, realism was correlated with the use of Facebook

XSL•FO
RenderX

(r=0.210), Instagram (r=0.208), and Snapchat (r=0.185), but not YouTube (r=0.074, *P*=.07).

Hypothesis 2b predicted that social media use would be positively associated with exposure to e-cigarette messages on social media. For each social media platform, the frequency of exposure to e-cigarette messages was regressed on to the frequency of use of the platform while controlling for age, sex, race, and social media use motivations. Table 4 presents the results. Across the platforms, social media use was significantly associated with exposure to e-cigarette messages.

Hypothesis 2c predicted that exposure to e-cigarette messages on social media would be positively associated with attitude toward e-cigarette use, which was regressed onto age, sex, race, social media use motivations, frequencies of using social media, and frequencies of exposure to e-cigarette messages on social media. The results are presented in Table 5. Significant positive associations between e-cigarette message exposure on Instagram and Snapchat and attitude toward e-cigarette use were found. The more frequently adolescents saw e-cigarette messages on each of these platforms, the more positive were their attitudes toward e-cigarette use.

Serial mediation analyses using the PROCESS path-analysis macro developed by Hayes [48] were conducted to examine the overall predictions of Hypotheses 2a-c, such that the effects of social media use motivations on attitude toward e-cigarette would be mediated through frequency of social media use and exposure to e-cigarette information on each of the four social media platforms. A bootstrapping estimation approach with 5000 samples was used to test the indirect effects in each model [49].

Agency motivation increased attitude toward e-cigarette use through frequency of Facebook use, independent of exposure to e-cigarette information on Facebook (point estimate=0.02, 95% CI 0.001-0.049), but not through exposure to e-cigarette information on Facebook, independent of frequency of Facebook use (point estimate=–0.002, 95% CI –0.016 to 0.007), and not serially through frequency of Facebook use and exposure to e-cigarette information on Facebook (point estimate=0.004, 95% CI –0.006 to 0.018). There was no direct effect of agency motivation on attitude toward e-cigarette use (point estimate=–0.05, 95% CI –0.174 to 0.072).

Table 3. Associations between motivations and frequency of social media use among adolescents.

Motivation	Social media platform							
	Facebook		Instagram		Snapchat		YouTube	
	B	β	B	β	B	β	B	β
Agency	0.433	0.197[a]	–0.153	–0.060	–0.264	–0.100	–0.058	–0.031
Filter	0.478	0.177[a]	0.218	0.071	0.302	0.093	0.112	0.054
Self-express	–0.165	–0.075	0.680	0.258[b]	0.651	0.237[b]	0.010	0.000
Social learning	–0.190	–0.072	0.515	0.178[a]	0.518	0.171[c]	0.082	0.044
Social comparison	0.121	0.049	–0.184	–0.069	–0.311	–0.111	0.381	0.197[a]
Realism	0.138	0.054	–0.019	–0.004	0.042	0.014	–0.164	–0.070

[a]*P*<.01.

[b]*P*<.001.

[c]*P*<.05.

Table 4. Associations between social media use frequency and e-cigarette message exposure frequency among adolescents.

Social media platform	B	Standard error	β	*P* value
Facebook	0.165	0.013	0.471	<.001
Instagram	0.163	0.014	0.459	<.001
Snapchat	0.174	0.013	0.487	<.001
YouTube	0.142	0.021	0.273	<.001

Table 5. Associations between exposure to e-cigarette messages on social media and e-cigarette use attitude among adolescents.

Social media platform	B	Standard error	β	*P* value
Facebook	–0.082	0.067	–0.071	.23
Instagram	0.214	0.076	0.217	.005
Snapchat	0.154	0.069	0.163	.03
YouTube	–0.079	0.056	–0.083	.15

XSL•FO
RenderX

Filter motivation increased attitude toward e-cigarette use through frequency of Facebook use independent of exposure to e-cigarette information on Facebook (point estimate=0.02, 95% CI 0.0002-0.055), not through exposure to e-cigarette information on Facebook, independent of frequency of Facebook use (point estimate=0.003, 95% CI –0.007 to 0.016), and not serially through frequency of Facebook use and exposure to e-cigarette information on Facebook (point estimate=0.005, 95% CI –0.007 to 0.019). There was no direct effect of filter motivation on attitude toward e-cigarette use (point estimate =–0.01, 95% CI –0.154 to 0.134).

Self-expression motivation increased attitude toward e-cigarette use serially through frequency of Instagram use and exposure to e-cigarette information on Instagram (point estimate=0.03, 95% CI 0.009 to 0.048), not through frequency of Instagram use, independent of exposure to e-cigarette information on Instagram (point estimate=–0.01, 95% CI –0.039 to 0.011), or through exposure to e-cigarette information on Instagram, independent of frequency of Instagram use (point estimate=–0.01, 95% CI –0.039 to 0.024). There was no direct effect of self-expression motivation on attitude toward e-cigarette use (point estimate=0.10, 95% CI –0.019 to 0.222; Figure 2).

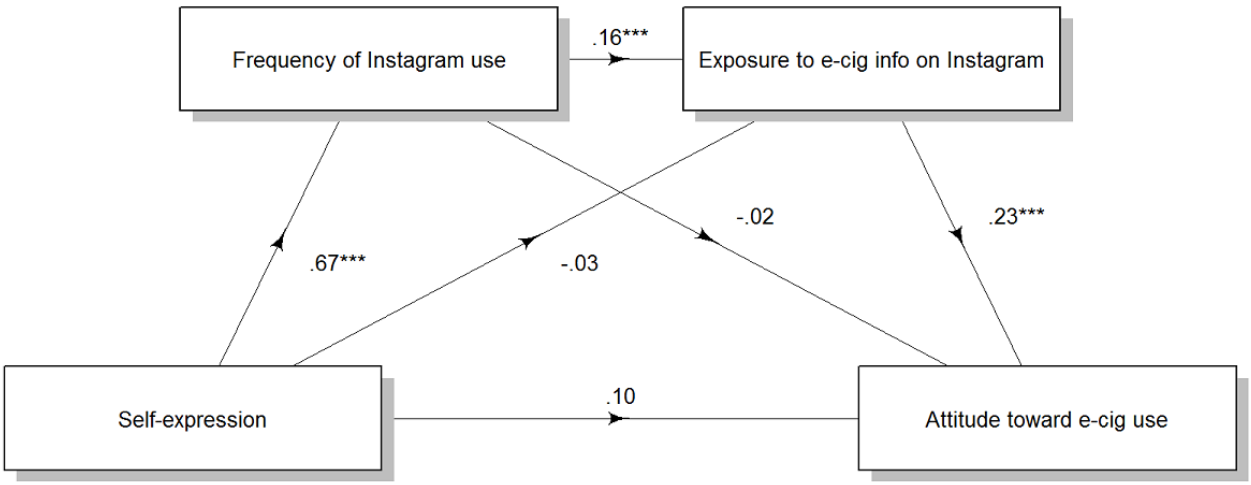
Similarly, self-expression motivation increased attitude toward e-cigarette use serially through frequency of Snapchat use and exposure to e-cigarette information on Snapchat (point estimate=0.03, 95% CI 0.009 to 0.049), not through frequency of Snapchat use, independent of exposure to e-cigarette information on Snapchat (point estimate=–0.002, 95% CI –0.026 to 0.019), and not through exposure to e-cigarette information on Instagram, independent of frequency of Snapchat use (point estimate=–0.007, 95% CI –0.023 to 0.035). There was no direct effect of self-expression motivation on attitude toward e-cigarette use (point estimate=0.08, 95% CI –0.041 to 0.199; Figure 3).

Social learning motivation increased attitude toward e-cigarette use serially through frequency of Instagram use and exposure to e-cigarette information on Instagram (point estimate=0.02, 95% CI 0.003-0.040), through exposure to e-cigarette information on Instagram, independent of frequency of Instagram use (point estimate=0.04, 95% CI 0.012 to 0.085), but not through frequency of Instagram use, independent of exposure to e-cigarette information on Instagram (point estimate=–0.008, 95% CI –0.031 to 0.009). There was no direct effect of social learning motivation on attitude toward e-cigarette use (point estimate=–0.03, 95% CI –0.175 to 0.111; Figure 4).

Similarly, social learning motivation increased attitude toward e-cigarette use serially through frequency of Snapchat use and exposure to e-cigarette information on Snapchat (point estimate=0.02, 95% CI 0.003-0.040), through exposure to e-cigarette information on Snapchat, independent of frequency of Snapchat use (point estimate=0.04, 95% CI 0.010-0.083), but not through frequency of Instagram use, independent of exposure to e-cigarette information on Instagram (point estimate=–0.002, 95% CI –0.021 to 0.016). There was no direct effect of social learning motivation on attitude toward e-cigarette use (point estimate=–0.04, 95% CI –0.183 to 0.103; Figure 5).

Lastly, the serial mediation from social comparison motivation to attitude toward e-cigarette use through frequency of YouTube use and exposure to e-cigarette information on YouTube was not significant. None of the paths in this model were significant.

Figure 2. Effects of self-expression motivation on attitude toward e-cigarette use through frequency of Instagram use and exposure to e-cigarette messages on Instagram. Numbers are unstandardized regression coefficients. **P*<.05, ***P*<.01, ****P*<.001. e-cig/e-cigarette: electronic cigarette.



XSL•FO
RenderX

Figure 3. Effects of self-expression motivation on attitude toward e-cigarette use through frequency of Snapchat use and exposure to e-cigarette messages on Snapchat use. Numbers are unstandardized regression coefficients. *P<.05, **P<.01, ***P<.001. e-cig/e-cigarette: electronic cigarette.

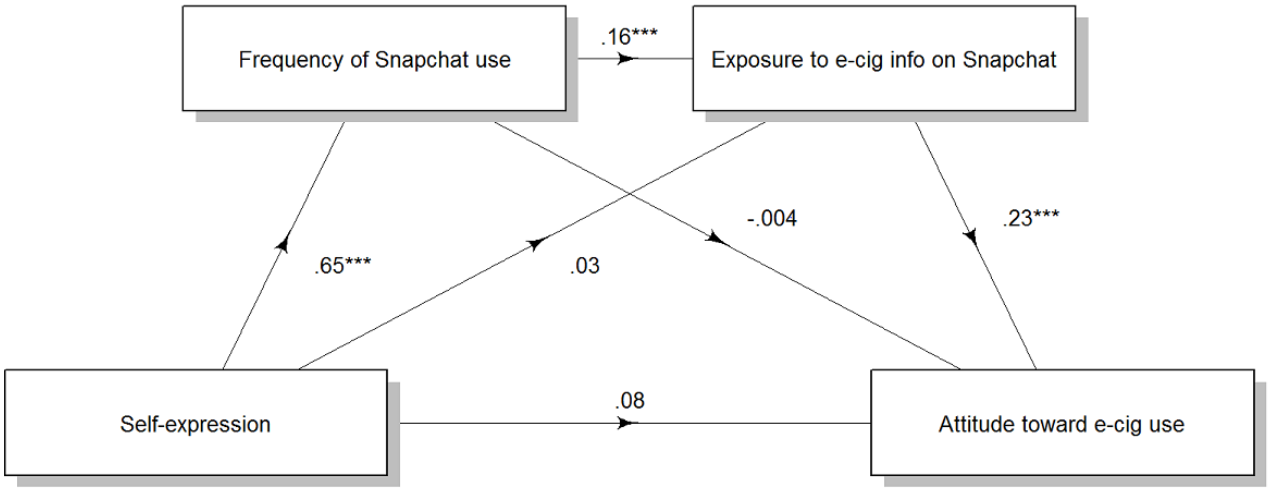


Figure 4. Effects of social learning motivation on attitude toward e-cigarette use through frequency of Snapchat use and exposure to e-cigarette messages on Snapchat. Numbers are unstandardized regression coefficients. *P<.05, **P<.01, ***P<.001. e-cig/e-cigarette: electronic cigarette.

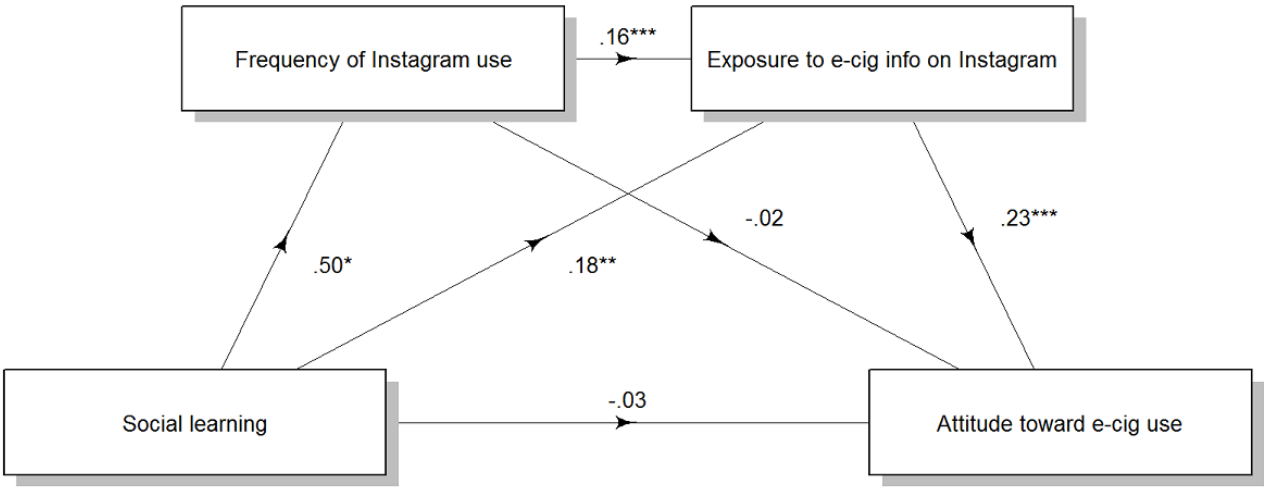
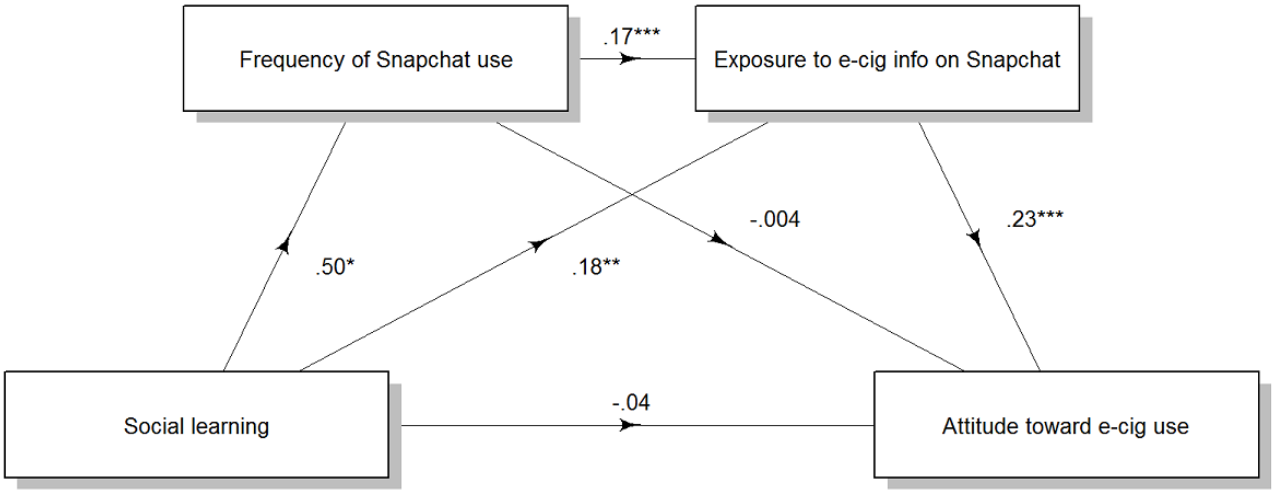


Figure 5. Effects of social learning motivation on attitude toward e-cigarette use through frequency of Instagram use and exposure to e-cigarette messages on Instagram use. Numbers are unstandardized regression coefficients. *P<.05, **P<.01, ***P<.001. e-cig/e-cigarette: electronic cigarette.



XSL•FO
RenderX

Hypothesis 3

Hypothesis 3a predicted that effects of social media–based exposure to e-cigarette messages will be differentially moderated by different social media use motivations such that high realism motivation will amplify the effect of the exposure, while low realism motivation will attenuate the effect. Moderation analyses using PROCESS macro with the 5000 bootstrapping estimation approach were conducted to examine the hypotheses [48,49].

The results showed that perceived realism interacted with exposure to e-cigarette messages on Facebook in influencing attitude toward e-cigarette use ($F_{1,562}$=4.31, P=.04). Further inspection indicated that exposure to e-cigarette information on Facebook is related to attitude among adolescents with high perceived realism (point estimate=0.17, P=.001), not for those with moderate or low perceived realism (Figure 6).

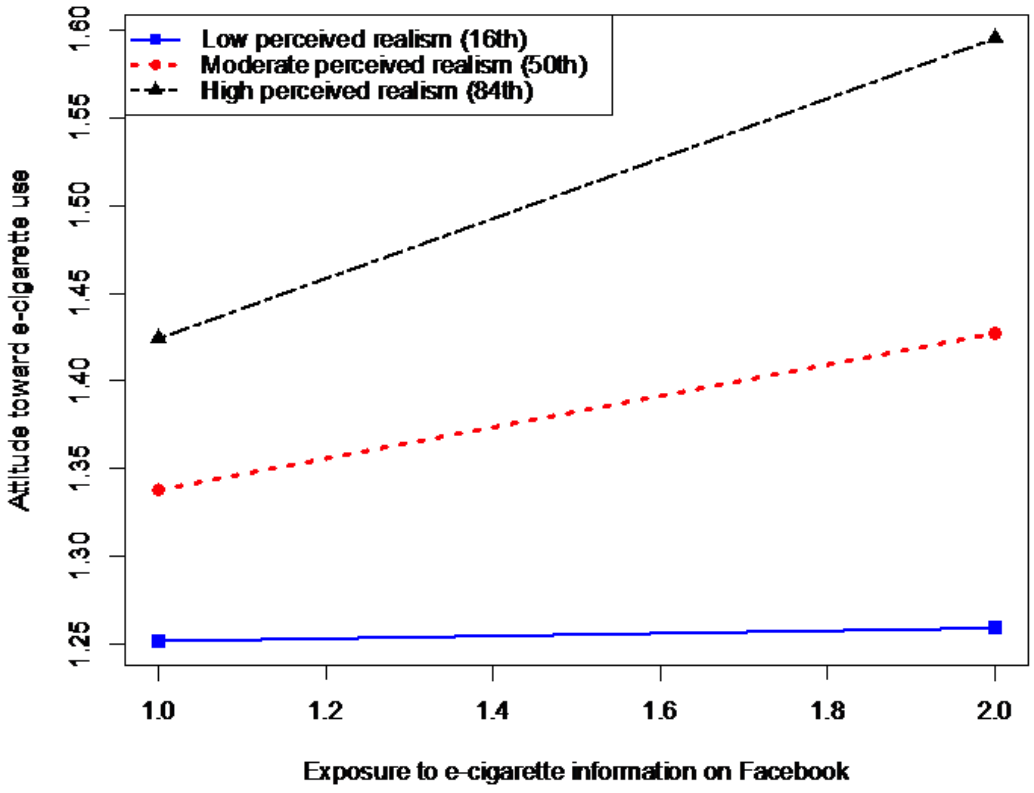
Perceived realism also moderated the relationship between exposure to e-cigarette messages on Instagram and the attitude toward e-cigarette use ($F_{1,568}$=15.88, P<.001). Further inspection indicated that exposure on Instagram was related to attitude among adolescents with moderate (point estimate=0.17, P<.001) and high perceived realism (point estimate=0.32, P<.001), not for those with low perceived realism (Figure 7).

Similarly, perceived realism moderated the relationship between exposure to e-cigarette use on Snapchat and attitude toward e-cigarette use ($F_{1,566}$=12.35, P<.001). This effect was observed only for adolescents with moderate (point estimate=.20, P<.001) and high realism (point estimate=0.32, P<.001), not for those with low realism (Figure 8).

Finally, another significant moderating effect was observed between perceived realism on exposure to e-cigarette messages on YouTube and the attitude toward e-cigarette use ($F_{1,569}$=13.90, P<.001). The effect was observed only for adolescents with moderate (point estimate=0.09, P<.05) and high perceived realism (point estimate=0.23, P<.001), not for those with low perceived realism (Figure 9).

Hypotheses 3b-d predicted three-way interactions among exposure, social motivations (ie, social learning, social comparison, and filter), and social norm. A significant three-way interaction among exposure on Instagram, social learning motivation, and social norm was found. As predicted, the moderating effect of social learning motivation on the relationship between exposure to e-cigarette information on Instagram and attitude toward e-cigarette use was found to depend on social norm (ie, the number of friends who use e-cigarettes; $F_{1,564}$=5.14, P=.02). The three-way interaction effect was significant for those who had more than a few friends who used e-cigarettes but not for those without friends who used e-cigarettes. This indicates that for adolescents who had greater e-cigarette use norms in their networks, the higher the social learning motivation, the more positive their attitude toward e-cigarette use when they were exposed to more e-cigarette information on Instagram. The three-way interaction is depicted in Figure 10. Hypothesis 3b concerning social learning motivation was supported, while Hypothesis 3c-d regarding social comparison and filter motivations was not supported.

Figure 6. Moderating effects of perceived realism and exposure to e-cigarette information on Facebook on attitude toward e-cigarette use. e-cigarette: electronic cigarette.



XSL•FO
RenderX

Figure 7. Moderating effects of perceived realism and exposure to e-cigarette information on Instagram on attitude toward e-cigarette use. e-cigarette: electronic cigarette.

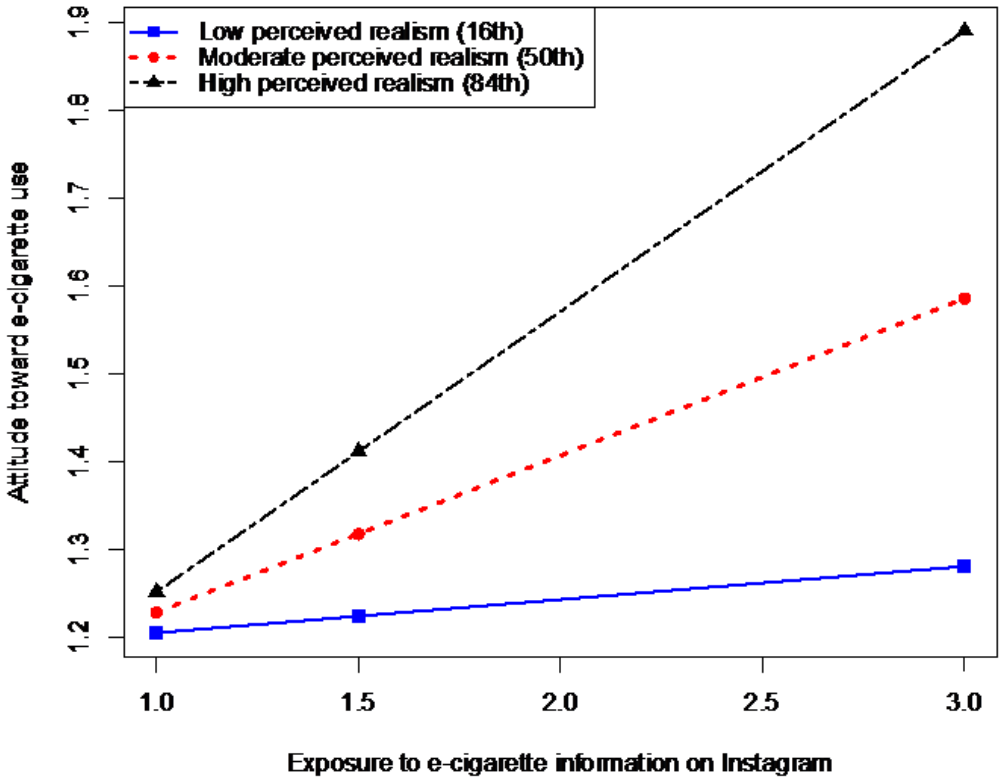
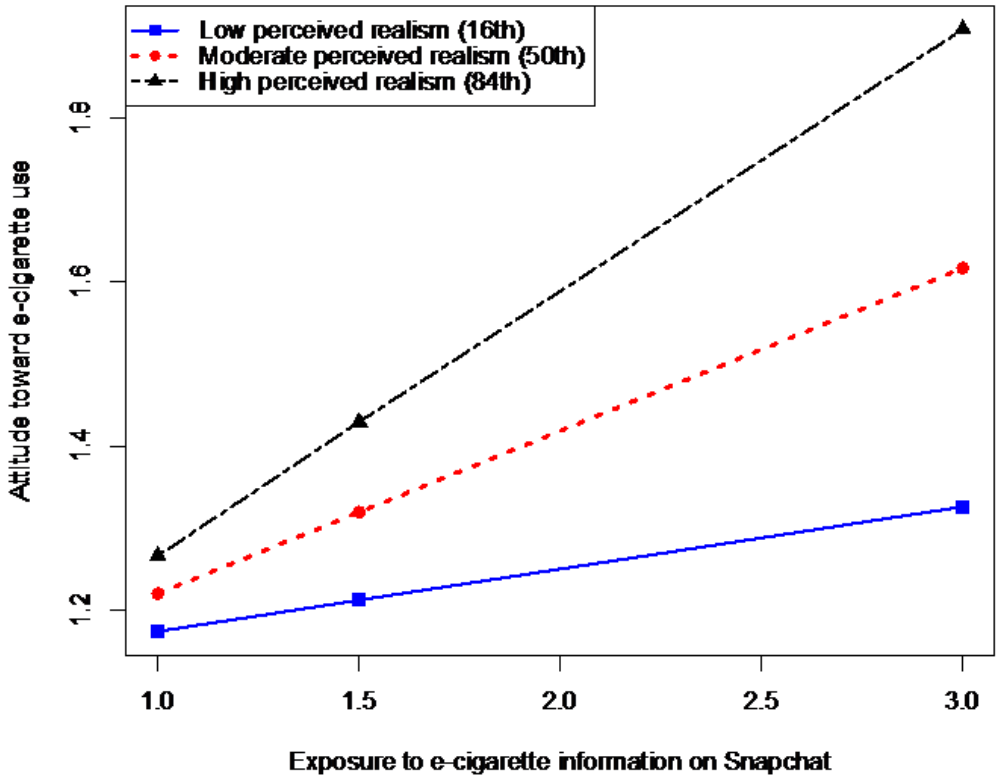


Figure 8. Moderating effects of perceived realism and exposure to e-cigarette information on Snapchat on attitude toward e-cigarette use. e-cigarette: electronic cigarette.



XSL•FO
RenderX

Figure 9. Moderating effects of perceived realism and exposure to e-cigarette information on YouTube on attitude toward e-cigarette use. e-cigarette: electronic cigarette.

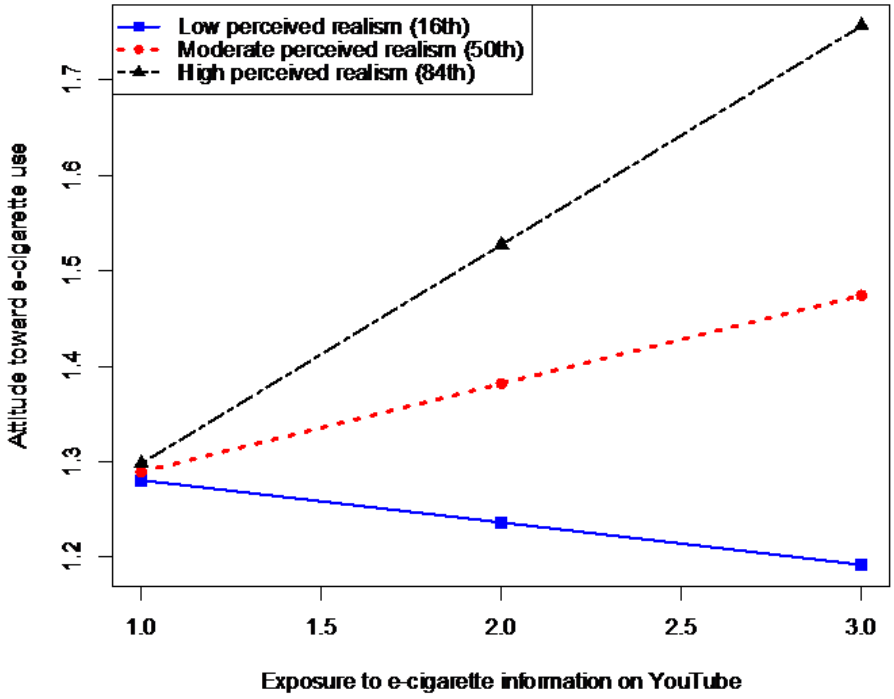
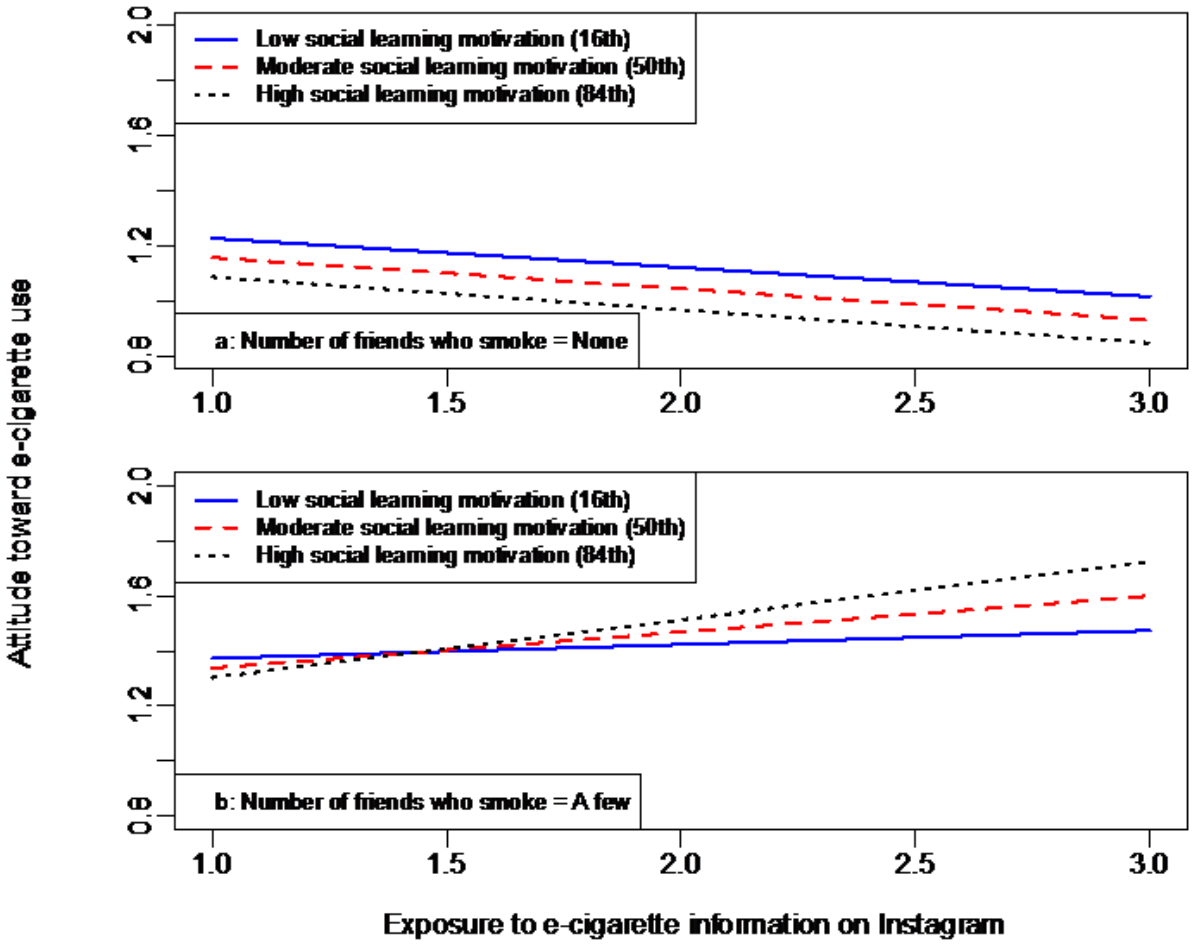


Figure 10. Effects of social norm on the moderating effects of social learning motivation on the relationship between exposure to e-cigarette information on Instagram and attitude toward e-cigarette use. e-cigarette: electronic cigarette.



XSL•FO
RenderX

Discussion

Principal Findings

This study aimed at investigating the mechanisms of the effects of social media use on attitude toward e-cigarette use among adolescents. To this end, we first investigated motivations of social media use. Next, we examined the roles of motivations in the process of effects of social media use and e-cigarette message exposure on e-cigarette use attitude. Collectively, the results contribute to extant theoretical and practical knowledge about the mechanisms of social media effects on risk exposure and risk attitude.

In a survey of a national sample of 14- to 17-year-old adolescents, we found that their motivations for social media use included agency, self-expression, realism, social learning, social comparison, and filter (Hypothesis 1). To our knowledge, this is a first validated measure of social media use motivation that is not platform-specific, uses a mixed attributes approach, and recognizes the unique functions of Web 2.0. Furthermore, other results of this study demonstrate the utility of the scale and its components in understanding various social media use and their effects on risk exposure and attitude.

The six motivations predicted individual social media platform use differentially (Hypothesis 2a). Agency and filter motivations were associated with Facebook use, whereas self-expression and social learning were associated with Instagram and Snapchat use. Dissimilar to other platforms, YouTube use was associated with social comparison motivation.

The parsimoniousness of our social media use motivation scale can facilitate research and practice, as it provides an understanding of differential motivations attached to divergent social media platforms and thereby assisting efforts to examine their effects. Practically, as social media use motivations indirectly predicted attitude toward e-cigarette use through frequency of social media use and/or exposure to e-cigarette messages on social media (Hypothesis 2a-c), understanding these motivations and addressing them will be essential to future prevention efforts. Of note, the influences of self-expression and social learning motivations on attitude toward e-cigarette use were greater than those of other motivations, indicating the importance of addressing these two motivations.

Across the four popular social media platforms among adolescents, frequency of use significantly predicted frequency of exposure to e-cigarette messages (Hypothesis 2b). Although not all exposure predicted positive attitude toward e-cigarette use (Hypothesis 2c), the positive associations between frequencies of social media use and e-cigarette message exposure suggest the importance of addressing social media–based exposure to e-cigarette messages.

More frequent exposure to e-cigarette messages on Instagram and Snapchat predicted more positive attitude toward e-cigarette use among adolescents (Hypothesis 2c), suggesting that these two social media may require greater attention than other platforms from prevention efforts.

The role of YouTube observed in this study is noteworthy. Although adolescents' volume of YouTube use was larger (Table 2), the association between exposure and attitude was not significant on YouTube (Hypothesis 2c). This may be because YouTube is a less networked platform than other social media such as Instagram and Snapchat. Exposure to e-cigarette messages on Facebook was lower than that on other social media, perhaps reflecting lower usage of the platform itself. This low usage of Facebook among adolescents found in this study is consistent with the latest report from Pew Research Center [4].

Notably, perceived realism significantly moderated the effects of social media–based e-cigarette message exposure on attitude toward e-cigarette use (Hypothesis 3a). When perceived realism was high, the exposure effect on attitude was amplified, but when perceived realism was low, the exposure effect was mitigated. This pattern was consistent across the four platforms of Facebook, Instagram, Snapchat, and YouTube (Figure 2). These findings indicate that social media realism judgment can be an important variable to include in future intervention efforts to change attitude and prevent e-cigarette use among adolescents.

Importantly, the perceived realism scale of this study captures the unique aspect of social media content. The social media realism scale of this study taps into the user-generated content and how it reflects peer users' true selves. In contrast, traditional mass media content comprises media professional–generated content. This difference can constitute an important conceptual distinction between (mass) media literacy [52] and social media literacy. Efforts to curb the harmful effects of social media on adolescents should be cognizant of this participative and user-generated aspect of the social media world and address the distinctiveness in the conceptualization and development of interventions.

The posited networked nature of "social" in social media was further demonstrated by the finding of a three-way interaction. Social norm moderated the effect of social learning motivation on the relationship between exposure and attitude (Hypothesis 3b). Among adolescents who did not have friends who used e-cigarette, the effect of social learning motivation was mitigated, but among those with friends who used e-cigarettes, the effect was amplified. These results suggest that research on the uses and effects of social media should consider the interpersonal and social contexts of social media use as well, incorporating both the online and offline networks, as adolescents' risk exposure and risk behavior may occur at this interface. Likewise, future e-cigarette control efforts should address the online and offline social environments of adolescents as they interface.

The same pattern of three-way interaction, however, was not found for other social motivations including comparison and filter (Hypothesis 3c-d), suggesting that their roles may be more complex than those of social learning motivation. As these motivations can play a role in the effects of social media use, future research should continue to examine these motivations and their interplay with the composition, structure, and size of social networks of adolescents.

XSL•FO
RenderX

Limitations and Suggestions for Future Research

Although the outcome variable of this study was attitude rather than behavior, meta-analyses have consistently demonstrated that attitude predicts behavior [53,54]. Moreover, as the occurrence of behavior can be contingent upon various structural factors that facilitate or hinder the behavior [55], this study's focus on attitude provides an up-close look at the relationships among motivations, uses, and effects of social media.

Future research could increase the number of items in some of the dimensions (eg, social comparison and filter; Multimedia Appendix 1) and test the scale with different populations to clarify the factor structure. Research may also benefit by scrutinizing the properties and functions of some of the dimensions of the scale, including social comparison and filter dimensions, as the posited three-way interaction among these dimensions, exposure, and social norm was not observed. Finally, research should continue to investigate and identify the core attributes, features, or affordances that cut across existing and emergent social media platforms, to better understand their uses and effects.

Implications for Theory and Practice

This study fills a gap in the literature on social media effects on e-cigarette use through its investigation of the mechanisms of social media effects on risk behavior likelihood among adolescents. Through the investigation, it identified motivations, mediators, and moderators of social media effects. The results show not only how the motivations and uses impact adolescents' attitude toward e-cigarette use, but also how the harmful effects could be mitigated. These findings inform the theory of social media effects and intervention for preventing harmful social media effects.

This study provides specific practical implications for future intervention efforts. It suggests that understanding and addressing the motivations associated with social media use are important, as it found that differential motivations have differential impact in the process of social media effects on e-cigarette use. Furthermore, this study shows that understanding and addressing self-expression and social learning motivations will be especially important, as these motivations exerted the strongest influence on frequency of social media use, exposure to e-cigarette messages, and attitude toward e-cigarette use. For example, providing youth with creative ways to participate in health campaigns could be a way to channel their self-expression motivation. A recent study found that production of digital counter messages by youth helped reduce risk behavior among them [56].

Instagram and Snapchat emerged as two of the more consequential social media platforms for adolescents. The results of this study show that the greater impact of these two platforms may stem from their capacities to satisfy the self-expression and social learning motivations of adolescents. As noted above, channeling these motivations to healthy directions will be important in future research and action for e-cigarette control. Instagram, especially, is noteworthy, as its visual focus may accommodate the visual nature and images of e-cigarette use behavior and their diffusion on social media and networks. Coupled with its significant impact on adolescents found in this study, the visual impact of Instagram merits further research. Recent research found that Instagram serves self-expressive and social engagement functions and that differential risk beliefs and emotions activate differential engagement and social support [57].

A new construct—social media realism—emerged as a significant moderator of social media effects. Social media realism consistently moderated the effects of exposure to e-cigarette message on e-cigarette attitude across adolescents' four most popular social media platforms. These results suggest that correcting social media realism is critical to addressing the harmful effects of social media. In light of its significant role identified in this study, future research should engage in in-depth investigation of social media realism, as this construct differs from past mass media–based conceptualizations and offers important directions for future interventions. This effort could be central to advancing the conceptual knowledge basis of social media literacy.

The three-way interaction among exposure, social learning motivation, and social norm indicates that efforts to reduce the effects of social media e-cigarette may not be fully efficacious if they do not take into account the social contexts of adolescents. As the moderating effect of social learning motivation is contingent upon the social normative environment of adolescents, intervention efforts should address both the online and offline social contexts of adolescent lives and risk vulnerability.

Conclusions

In summary, this study contributes to conceptual knowledge about the process of social media effects on risk behavior attitudes. The motivations of social media use and their mediating and moderating roles identified in this study will inform future research.

Acknowledgments

This study was supported in part by a grant (R01CA176196) from the National Cancer Institute of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Social Media Use Motivation Scale.

[PNG File, 391KB - jmir_v21i6e14303_app1.png]

References

1. Hendriks H, Van den Putte B, Gebhardt WA, Moreno MA. Social Drinking on Social Media: Content Analysis of the Social Aspects of Alcohol-Related Posts on Facebook and Instagram. J Med Internet Res 2018 Jun 22;20(6):e226 [FREE Full text] [doi: 10.2196/jmir.9355] [Medline: 29934290]
2. Moreno MA, Gower AD, Jenkins MC, Kerr B, Gritton J. Marijuana promotions on social media: adolescents' views on prevention strategies. Subst Abuse Treat Prev Policy 2018 Dec 02;13(1):23 [FREE Full text] [doi: 10.1186/s13011-018-0152-7] [Medline: 29961426]
3. Best P, Manktelow R, Taylor B. Online communication, social media and adolescent wellbeing: A systematic narrative review. Children and Youth Services Review 2014 Jun;41:27-36. [doi: 10.1016/j.childyouth.2014.03.001]
4. Pew Research Center. Teens, Social Media and Technology 2018. URL: https://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/ [accessed 2019-06-01] [WebCite Cache ID 77RX8qvMP]
5. Hajek P, Phillips-Waller A, Przulj D, Pesola F, Myers Smith K, Bisal N, et al. A Randomized Trial of E-Cigarettes versus Nicotine-Replacement Therapy. N Engl J Med 2019 Dec 14;380(7):629-637. [doi: 10.1056/NEJMoa1808779] [Medline: 30699054]
6. Goniewicz ML, Smith DM, Edwards KC, Blount BC, Caldwell KL, Feng J, et al. Comparison of Nicotine and Toxicant Exposure in Users of Electronic Cigarettes and Combustible Cigarettes. JAMA Netw Open 2018 Dec 07;1(8):e185937 [FREE Full text] [doi: 10.1001/jamanetworkopen.2018.5937] [Medline: 30646298]
7. England L, Bunnell R, Pechacek T, Tong V, McAfee T. Nicotine and the Developing Human: A Neglected Element in the Electronic Cigarette Debate. Am J Prev Med 2015 Aug;49(2):286-293 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.015] [Medline: 25794473]
8. The Health Consequences of Smoking: Nicotine Addiction. A Report of the Surgeon General. Atlanta, GA: Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 1988. URL: https://profiles.nlm.nih.gov/ps/access/nnbbzd.pdf [accessed 2019-06-11]
9. Jensen RP, Luo W, Pankow JF, Strongin RM, Peyton DH. Hidden formaldehyde in e-cigarette aerosols. N Engl J Med 2015 Jan 22;372(4):392-394. [doi: 10.1056/NEJMc1413069] [Medline: 25607446]
10. Surgeon General's Advisory on E-cigarette Use Among Youth. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2018. URL: https://e-cigarettes.surgeongeneral.gov/documents/surgeon-generals-advisory-on-e-cigarette-use-among-youth-2018.pdf [accessed 2019-06-11] [WebCite Cache ID 77UWwctlK]
11. Katz E, Blumler J, Gurevitch M. Uses and gratifications research. Pub Opinion Quart 1973;37(4):509-523. [doi: 10.1086/268109]
12. Rubin AM. An Examination of Television Viewing Motivations. Commun Res 1981;8(2):141-165. [doi: 10.1177/009365028100800201] [Medline: 22973420]
13. Rubin AM. The uses-and-gratifications perspective of media effects. In: Media Effects: Advances in Theory and Research. Mahwah, NJ: Erlbaum; 2009:161-184.
14. So J. Uses, Gratifications, and Beyond: Toward a Model of Motivated Media Exposure and Its Effects on Risk Perception. Communication Theory 2012;22(2):116-137. [doi: 10.1111/j.1468-2885.2012.01400.x]
15. Raacke J, Bonds-Raacke J. MySpace and Facebook: applying the uses and gratifications theory to exploring friend-networking sites. Cyberpsychol Behav 2008 Apr;11(2):169-174. [doi: 10.1089/cpb.2007.0056] [Medline: 18422409]
16. Chen GM. Tweet this: A uses and gratifications perspective on how active Twitter use gratifies a need to connect with others. Computers in Human Behavior 2011 Mar;27(2):755-762. [doi: 10.1016/j.chb.2010.10.023]
17. Sheldon P, Bryant K. Instagram: Motives for its use and relationship to narcissism and contextual age. Computers in Human Behavior 2016 May;58:89-97. [doi: 10.1016/j.chb.2015.12.059]
18. Haridakis P, Hanson G. Social Interaction and Co-Viewing With YouTube: Blending Mass Communication Reception and Social Connection. Journal of Broadcasting & Electronic Media 2009 May 27;53(2):317-335. [doi: 10.1080/08838150902908270]
19. Utz S, Muscanell N, Khalid C. Snapchat Elicits More Jealousy than Facebook: A Comparison of Snapchat and Facebook Use. Cyberpsychology, Behavior, and Social Networking 2015 Mar;18(3):141-146. [doi: 10.1089/cyber.2014.0479]
20. Eveland W. A Mix of Attributes approach to the study of media effects and new communication technologies. J Commun 2003;53(3):395-410. [doi: 10.1111/j.1460-2466.2003.tb02598.x]
21. Sundar SS, Limperos AM. Uses and Grats 2.0: New Gratifications for New Media. Journal of Broadcasting & Electronic Media 2013 Oct;57(4):504-525. [doi: 10.1080/08838151.2013.845827]

XSL•FO
RenderX

22. Shao G. Understanding the appeal of user ‐ generated media: a uses and gratification perspective. Internet Research 2009 Jan 30;19(1):7-25. [doi: 10.1108/10662240910927795]
23. Walther JB, Jang J. Communication Processes in Participatory Websites. J Comput-Mediat Comm 2012 Oct 10;18(1):2-15. [doi: 10.1111/j.1083-6101.2012.01592.x]
24. Moreno MA, Kota R, Schoohs S, Whitehill JM. The Facebook influence model: a concept mapping approach. Cyberpsychol Behav Soc Netw 2013 Jul;16(7):504-511 [FREE Full text] [doi: 10.1089/cyber.2013.0025] [Medline: 23621717]
25. Östman J. Information, expression, participation: How involvement in user- generated content relates to democratic engagement among young people. New Media & Society 2012 Mar 20;14(6):1004-1021. [doi: 10.1177/1461444812438212]
26. Ekström M, Östman J. Information, Interaction, and Creative Production. Communication Research 2013 Feb 21;42(6):796-818. [doi: 10.1177/0093650213476295]
27. Turkle S. Life on the Screen: Identity in the Age of the Internet. New York: Simon & Schuster; 1995.
28. Cho H, Shen L, Wilson K. Perceived Realism. Communication Research 2014;41(6):828-851. [doi: 10.1177/0093650212450585]
29. Rubin AM, Perse EM. Audience Activity and Soap Opera Involvement A Uses and Effects Investigation. Human Comm Res 1987 Dec;14(2):246-268. [doi: 10.1111/j.1468-2958.1987.tb00129.x]
30. Quan-Haase A, Young AL. Uses and Gratifications of Social Media: A Comparison of Facebook and Instant Messaging. Bulletin of Science, Technology & Society 2010 Sep 14;30(5):350-361. [doi: 10.1177/0270467610380009]
31. Yee N, Bailenson JN, Urbanek M, Chang F, Merget D. The Unbearable Likeness of Being Digital: The Persistence of Nonverbal Social Norms in Online Virtual Environments. CyberPsychology & Behavior 2007 Feb;10(1):115-121. [doi: 10.1089/cpb.2006.9984]
32. Vogel EA, Rose JP, Roberts LR, Eckles K. Social comparison, social media, and self-esteem. Psychology of Popular Media Culture 2014;3(4):206-222. [doi: 10.1037/ppm0000047]
33. Yang C, Holden SM, Carter MDK. Social Media Social Comparison of Ability (but not Opinion) Predicts Lower Identity Clarity: Identity Processing Style as a Mediator. J Youth Adolescence 2018 Jan 11;47(10):2114-2128. [doi: 10.1007/s10964-017-0801-6]
34. Pariser E. The Filter Bubble: What the Internet Is Hiding from You. London, UK: The Penguin Press; 2011.
35. Flaxman S, Goel S, Rao JM. Filter Bubbles, Echo Chambers, and Online News Consumption. PUBOPQ 2016 Mar 22;80(S1):298-320. [doi: 10.1093/poq/nfw006]
36. Sunstein C. Republic.com 2.0. Princeton, NJ: Princeton University Press; 2007.
37. Messing S, Westwood SJ. Selective Exposure in the Age of Social Media. Communication Research 2012 Dec 31;41(8):1042-1063. [doi: 10.1177/0093650212466406]
38. Gibson J. The Ecological Approach to Visual Perception. Boston, MA: Houghton, Mifflin and Company; 1979.
39. Norman D. The Design of Everyday Things. New York, NY: Doubleday; 1990.
40. Fox J, McEwan B. Distinguishing technologies for social interaction: The perceived social affordances of communication channels scale. Communication Monographs 2017 Jun 28;84(3):298-318. [doi: 10.1080/03637751.2017.1332418]
41. Cole-Lewis H, Pugatch J, Sanders A, Varghese A, Posada S, Yun C, et al. Social Listening: A Content Analysis of E-Cigarette Discussions on Twitter. J Med Internet Res 2015 Oct 27;17(10):e243 [FREE Full text] [doi: 10.2196/jmir.4969] [Medline: 26508089]
42. McCausland K, Maycock B, Leaver T, Jancey J. The Messages Presented in Electronic Cigarette-Related Social Media Promotions and Discussion: Scoping Review. J Med Internet Res 2019 Feb 05;21(2):e11953 [FREE Full text] [doi: 10.2196/11953] [Medline: 30720440]
43. Huang J, Kornfield R, Emery SL. 100 Million Views of Electronic Cigarette YouTube Videos and Counting: Quantification, Content Evaluation, and Engagement Levels of Videos. J Med Internet Res 2016 Mar 18;18(3):e67 [FREE Full text] [doi: 10.2196/jmir.4265] [Medline: 26993213]
44. Pokhrel P, Fagan P, Herzog TA, Laestadius L, Buente W, Kawamoto CT, et al. Social media e-cigarette exposure and e-cigarette expectancies and use among young adults. Addictive Behaviors 2018 Mar;78:51-58. [doi: 10.1016/j.addbeh.2017.10.017]
45. Osgood C, Suci G, Tannenbaum P. The Measurement of Meaning. Urbana, IL: University of Illinois Press; 1957.
46. Fishbein M, Hall-Jamieson K, Zimmer E, von Haeften I, Nabi R. Avoiding the boomerang: testing the relative effectiveness of antidrug public service announcements before a national campaign. Am J Public Health 2002 Feb;92(2):238-245. [doi: 10.2105/ajph.92.2.238] [Medline: 11818299]
47. Hayes A. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. Second edition. New York, NY: Guilford Press; 2018.
48. Shrout PE, Bolger N. Mediation in experimental and nonexperimental studies: New procedures and recommendations. Psychological Methods 2002;7(4):422-445. [doi: 10.1037//1082-989X.7.4.422]
49. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychological Methods 1996;1(2):130-149. [doi: 10.1037/1082-989X.1.2.130]
50. Browne MW, Cudeck R. Alternative Ways of Assessing Model Fit. Sociological Methods & Research 2016 Jun 29;21(2):230-258. [doi: 10.1177/0049124192021002005]

XSL•FO
RenderX

51. Raftery AE. Bayesian Model Selection in Social Research. Sociological Methodology 1995;25:111. [doi: 10.2307/271063]
52. Jeong S, Cho H, Hwang Y. Media literacy interventions: a meta-analytic review. J Commun 2012;62(3):454-472 [FREE Full text] [doi: 10.1111/j.1460-2466.2012.01643.x] [Medline: 22736807]
53. Glasman LR, Albarracín D. Forming attitudes that predict future behavior: a meta-analysis of the attitude-behavior relation. Psychol Bull 2006 Sep;132(5):778-822 [FREE Full text] [doi: 10.1037/0033-2909.132.5.778] [Medline: 16910754]
54. Kim M, Hunter J. Attitude-behavior relations: A meta-analysis of attitudinal relevance and topic. J Commun 1993;43(1):101-142. [doi: 10.1111/j.1460-2466.1993.tb01251.x]
55. Fishbein M, Yzer MC. Using Theory to Design Effective Health Behavior Interventions. Commun Theory 2003 May;13(2):164-183. [doi: 10.1111/j.1468-2885.2003.tb00287.x]
56. Cho H, Yu B, Cannon J, Zhu Y. Efficacy of a media literacy intervention for indoor tanning prevention. J Health Commun 2018;23(7):643-651 [FREE Full text] [doi: 10.1080/10810730.2018.1500659] [Medline: 30058942]
57. Cho H, Silver N, Na K, Adams D, Luong KT, Song C. Visual Cancer Communication on Social Media: An Examination of Content and Effects of #Melanomasucks. J Med Internet Res 2018;20(9):e10501 [FREE Full text] [doi: 10.2196/10501] [Medline: 30185403]

Abbreviations

e-cigarette: electronic cigarette
GFI: goodness-of-fit index
CFI: comparative fit index
RMSEA: root mean square error of approximation
BIC: Bayesian information criterion

Edited by G Eysenbach; submitted 08.04.19; peer-reviewed by A Gower, R Hyzer; comments to author 02.05.19; revised version received 13.05.19; accepted 29.05.19; published 27.06.19.

Please cite as:
Cho H, Li W, Shen L, Cannon J
Mechanisms of Social Media Effects on Attitudes Toward E-Cigarette Use: Motivations, Mediators, and Moderators in a National Survey of Adolescents
J Med Internet Res 2019;21(6):e14303
URL: http://www.jmir.org/2019/6/e14303/
doi:10.2196/14303
PMID:31250830

©Hyunyi Cho, Wenbo Li, Lijiang Shen, Julie Cannon. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Monitoring Physical Activity Levels Using Twitter Data: Infodemiology Study

Sam Liu[1], PhD; Brian Chen[2], BSc; Alex Kuo[2], PhD

[1]School of Exercise Science, Physical and Health Education, University of Victoria, Victoria, BC, Canada

[2]Department of Health Information Science, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Sam Liu, PhD
School of Exercise Science, Physical and Health Education
University of Victoria
McKinnon Building 124
Victoria, BC, V8W 2Y2
Canada
Phone: 1 250 721 8392
Email: samliu@uvic.ca

Abstract

Background: Social media technology such as Twitter allows users to share their thoughts, feelings, and opinions online. The growing body of social media data is becoming a central part of infodemiology research as these data can be combined with other public health datasets (eg, physical activity levels) to provide real-time monitoring of psychological and behavior outcomes that inform health behaviors. Currently, it is unclear whether Twitter data can be used to monitor physical activity levels.

Objective: The aim of this study was to establish the feasibility of using Twitter data to monitor physical activity levels by assessing whether the frequency and sentiment of physical activity–related tweets were associated with physical activity levels across the United States.

Methods: Tweets were collected from Twitter's application programming interface (API) between January 10, 2017 and January 2, 2018. We used Twitter's *garden hose* method of collecting tweets, which provided a random sample of approximately 1% of all tweets with location metadata falling within the United States. Geotagged tweets were filtered. A list of physical activity–related hashtags was collected and used to further classify these geolocated tweets. Twitter data were merged with physical activity data collected as part of the Behavioral Risk Factor Surveillance System. Multiple linear regression models were fit to assess the relationship between physical activity–related tweets and physical activity levels by county while controlling for population and socioeconomic status measures.

Results: During the study period, 442,959,789 unique tweets were collected, of which 64,005,336 (14.44%) were geotagged with latitude and longitude coordinates. Aggregated data were obtained for a total of 3138 counties in the United States. The mean county-level percentage of physically active individuals was 74.05% (SD 5.2) and 75.30% (SD 4.96) after adjusting for age. The model showed that the percentage of physical activity–related tweets was significantly associated with physical activity levels (beta=.11; SE 0.2; P<.001) and age-adjusted physical activity (beta=.10; SE 0.20; P<.001) on a county level while adjusting for both Gini index and education level. However, the overall explained variance of the model was low (R^2=.11). The sentiment of the physical activity–related tweets was not a significant predictor of physical activity level and age-adjusted physical activity on a county level after including the Gini index and education level in the model (P>.05).

Conclusions: Social media data may be a valuable tool for public health organizations to monitor physical activity levels, as it can overcome the time lag in the reporting of physical activity epidemiology data faced by traditional research methods (eg, surveys and observational studies). Consequently, this tool may have the potential to help public health organizations better mobilize and target physical activity interventions.

(J Med Internet Res 2019;21(6):e12394) doi:10.2196/12394

KEYWORDS

physical activity; social media; internet; social media; Twitter messaging; population surveillance; public health

XSL•FO
RenderX

Introduction

Background

Physical inactivity is a modifiable risk factor for developing a widening variety of chronic conditions including cardiovascular diseases, hypertension, type 2 diabetes mellitus, colon cancers, osteoporosis, and depression [1-5]. Currently, many adults in the United States are physically inactive and do not meet the recommended amount of physical activity (150 min of moderate-intensity aerobic exercise per week) [6]. Furthermore, the prevalence of physical activity varies across geographic regions in the United States [7]. This lack of uniformity in the rate of physical activity in various geographic regions has become one of the top priorities for public health agencies—to collect population-level physical activity data. This epidemiology data can help identify groups and populations who are not engaged in regular physical activity and locations where these individuals live [8-10]. Local public health agencies can use this information to deploy appropriate resources to target health promotion efforts to improve physical activity levels in physically inactive regions. In fact, several studies have demonstrated the feasibility of using social media and internet-based interventions to promote physical activity on a population level [11-15]. Real-time epidemiology data of groups and locations of individuals who are not engaged in regular physical activity may further enhance public health agencies' capabilities to personalize and target their interventions.

Existing methods of using population-based survey studies to monitor physical activity need to be improved. There are several limitations in current methods of collecting physical activity data [10,16]. First, reporting physical activity survey data in the United States involves up to 2 to 3 years of lag time, whereas surveys themselves can be time-consuming and resource-intensive to conduct. Sparsity of data can be challenging for many surveys, as response rates may not vary uniformly by location, demographics, or population. Therefore, innovative research approaches are needed to supplement and improve the current state of physical activity monitoring.

Social media use has grown rapidly in the last decade, [17] and researchers have been examining ways to use *social data* to better understand and monitor public health problems in *real-time* [16,18]. This growing area of research has been called *infodemiology* or *infoveillance* studies [19,20]. Social media technology, such as Twitter, allows users to communicate with each other by sharing short messages. Users can share their thoughts, feelings, and opinions on these social media platforms and, as a result, social media data may be used to provide real-time monitoring of behavioral outcomes that inform health behaviors [21]. A unique aspect of social media data from Twitter is that the posts are public and geotagged and thus, all internet users, including health researchers, can readily access these data. In addition, unique to Twitter is the use of hashtags (#) that allow a user to highlight and allow other users to follow relevant topics of interest. Given their high level of use, these sites collect an enormous amount of data (eg, over 500 million tweets per day on Twitter) [21].

Recent infodemiology studies have reported that data from social media technologies can be combined with other biomedical datasets to help predict health outcomes [16,22-25]. The main approaches to analyze unstructured text data from Twitter include frequency of keyword occurrence (analysis of information prevalence and information occurrence ratio) and the sentiment of the tweets [10,19]. These approaches are not mutually exclusive and thus can be used together for monitoring physical activity. Information prevalence and information occurrence ratios measure the absolute or relative *frequency* of occurrences of a certain keyword. The amount of social data is constantly increasing; thus, normalized indicators (eg, information occurrence ratio) may be more meaningful than absolute figures on information prevalence [19]. Finally, sentiment analysis can determine whether an individual's attitude or perception toward a topic is positive, negative, or neutral. By applying these methods, researchers have shown that social data can be used to identify symptoms associated with psychological distress, anxiety, and depression [22] and identify infectious disease outbreaks, such as influenza transmission [26,27] and HIV outbreaks [24]. Previous studies have also reported that the frequency of physical activity–related tweets and the sentiment of the tweets are related to obesity rates [28], social disparity, and wellness indicators in US Metropolitan Statistical Areas (MSAs) [29,30]. Currently, it remains unclear whether these methods of analyzing physical activity–related tweets can be applied to monitor the physical activity level on a county level across the United States while controlling for socioeconomic inequality and education level.

Objectives

The aim of this study was to establish the feasibility of using Twitter data to monitor physical activity levels by assessing whether the frequency and sentiment of physical activity–related tweets were associated with physical activity levels in various counties across the United States.

Methods

Overview

Tweets (n=442,959,789) were collected from January 10, 2017, to January 2, 2018, using Twitter's application programming interface (API). The captured tweets represent an estimated 1% random selection of all tweets posted in a selected time frame. Only *geolocated* tweets with coordinates or within the bounding box defined by –162.354635, 18.756125, –53.755999, 73.893030 were retained for analysis. Additional processing was applied to filter out tweets with coordinates not originating from the United States, leaving a final sample of 64,005,336 tweets. To categorize tweets on a county level, a reverse-geocoding pipeline using cartographic boundary shapefiles from the US Census Bureau was created to assign a Federal Information Processing Standard code for each tweet.

Classifying Physical Activity–Related Tweets

A list of physical activity–related hashtags was compiled (see Multimedia Appendix 1) to identify tweets that might be related to exercise or physical activity. The hashtags were compiled using a combination of the most popular physical

XSL•FO
RenderX

activity–related hashtags and the guidelines for exercise testing published by the American College of Sports Medicine (ACSM) [31]. The ACSM guideline was used because it provided an extensive list of physical activity–related keywords that were well established, and this method has been used in previous research [28,31]. A tweet was classified as a physical activity–related tweet if it contained one or more physical activity–related keyword in the tweet's hashtags. Although previous studies have relied on dictionaries of exercise-related keywords (eg, from the Compendium of Physical Activities and ACSM guidelines for exercise testing) to classify tweets, using hashtags presents a couple of important advantages: they can be parsed as distinct entities from tweets and can represent more specific multi-word phrases [28,29]. As such, there is less risk of ambiguity with hashtags than with a dictionary or list of keywords (eg, *walk* and *surf* may have multiple meanings outside of physical activity, whereas *#30daysoffitness* is unlikely to). Previous research has also discussed the difficulty of this classification task, either electing to improve precision by imposing additional rules (eg, requiring additional context for commonly ambiguous terms such as *running*) on top of the basic dictionary word check list or choosing not to apply any additional filtering to avoid introducing additional biases into the sampling methodology [29]. Although using hashtags does not rectify the issue of curation bias, it does allow for far more specific matching against text than regular words, while also maintaining the simplicity of a simple list of items. This inherently trades off increased precision at the loss of recall or sensitivity but ensures that fewer unrelated tweets are passed to the sentiment analysis pipeline.

Sentiment Analysis

Using sentiment analysis techniques to study microblogging services such as Twitter is a rich and active area of study. Sentiment analysis assigns text documents polarities, labels such as *positive*, *negative*, and *neutral* that describe the writer's attitude as written. When applied to a topic, sentiment analysis may be used to predict or infer these attitudes based primarily on a collection of documents. This study utilized a sentiment analysis model created by Baziotis, Pelekis, and Doulkeridis for the 2017 International Workshop on Semantic Evaluation (SemEval) [32,33]. This model ranked first in Subtask A of Task 4 (*Sentiment Analysis in Twitter*) at SemEval 2017 and employs a bidirectional long short-term memory neural network with an attention mechanism [34].

Physical Activity Dataset

Physical activity levels and age-adjusted physical activity levels were extracted from the Behavioral Risk Factor Surveillance System (BRFSS) surveys, which provides county-level data of physical activity levels from the year 2014. The BRFSS is administered by the Centers for Disease Control and Prevention. As part of the survey, participants were asked to self-report leisure-time physical activity (eg, during the past month, other than your regular job, did you participate in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise?). Self-reported leisure-time physical inactivity was ascertained from answers of *no* to the questions. The BRFSS physical activity data were collected for adults aged 18 years and older, thus BRFSS also reported age-adjusted physical activity data based on the US standard population. Socioeconomic status measures, such as the Gini index, were collected from the American Community Survey.

Statistical Analysis

The frequency of physical activity tweets was tallied from each county and merged with physical activity levels from the BRFSS data, Gini index data, and percent of the county that received college education. The Gini index provides a standardized estimate of income inequality that may be used for comparisons between counties. Including the Gini index and education level is pertinent in the context of physical activity; these variables have been associated with levels of physical activity [35].

A bivariate Spearman correlation was used to determine the association between the number of physical activity–related tweets (including the number of positive, negative, and neutral tweets), the Gini index, education, and physical activity data. Multiple linear regression models were then applied to find the level of association between the proportion of physical activity–related tweets, sentiment of the tweets (ratio of positive to negative physical activity–related tweets), and physical activity data while controlling for the Gini index and education level on a county level. The relative performance of these models was compared. All analyses were performed using IBM SPSS 24.0 (IBM Corporation).

Results

Of the 442,959,789 unique tweets collected, 64,005,336 (14.44%) were geotagged. Of these, 234,678 (0.37%) were identified to be physical activity–related based on their hashtags. Los Angeles County (n=20,589; 8.77%), New York County (n=12,601; 5.37%), Miami–Dade County (n=7055; 3.01%), Harris County (n=6148; 2.62%), and Cook County (n=5738; 2.45%) were the 5 counties that sent the most geotagged physical activity–related tweets (Figure 1).

Aggregated data were obtained for a total of 3138 out of 3146 counties or county equivalents. The counties omitted for analysis lacked correlated Twitter data, physical activity data, or data on the socioeconomic indicators. The mean county-level percentage of individuals that are physically active was 74.05% (SD 5.2) and 75.30% (SD 4.96) after adjusting for age (Figure 2). Maps of the Gini index and education levels are displayed in Figures 3 and 4.

Our sentiment analysis showed that 7.31% (n=17,155) of the physical activity–related tweets identified were positive, 42.67% (n=100,137) were negative, and 50.02% (n=117,386) were neutral. The mean ratio between positive and negative was 0.20 (SD 0.336). Textbox 1 shows example tweets. On the basis of the correlation analysis, county-level physical activity and age-adjusted physical activity level showed a significant positive weak-to-moderate correlation with the percentage of physical activity–related tweets and the sentiment of physical activity–related tweets (Table 1).

XSL•FO
RenderX

Figure 1. Map of physical activity levels in the United States.

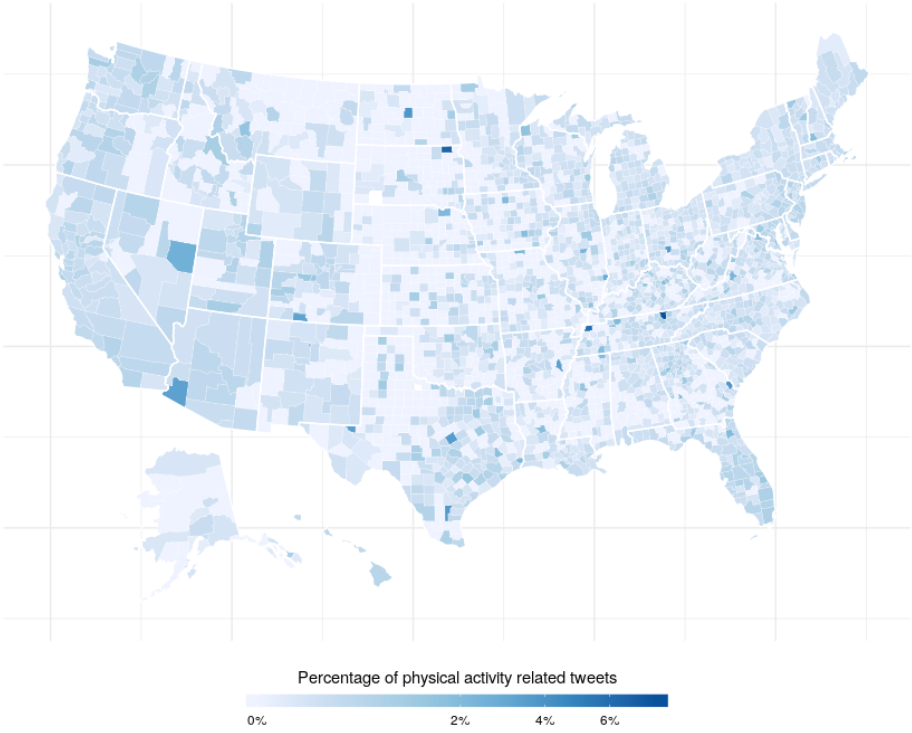
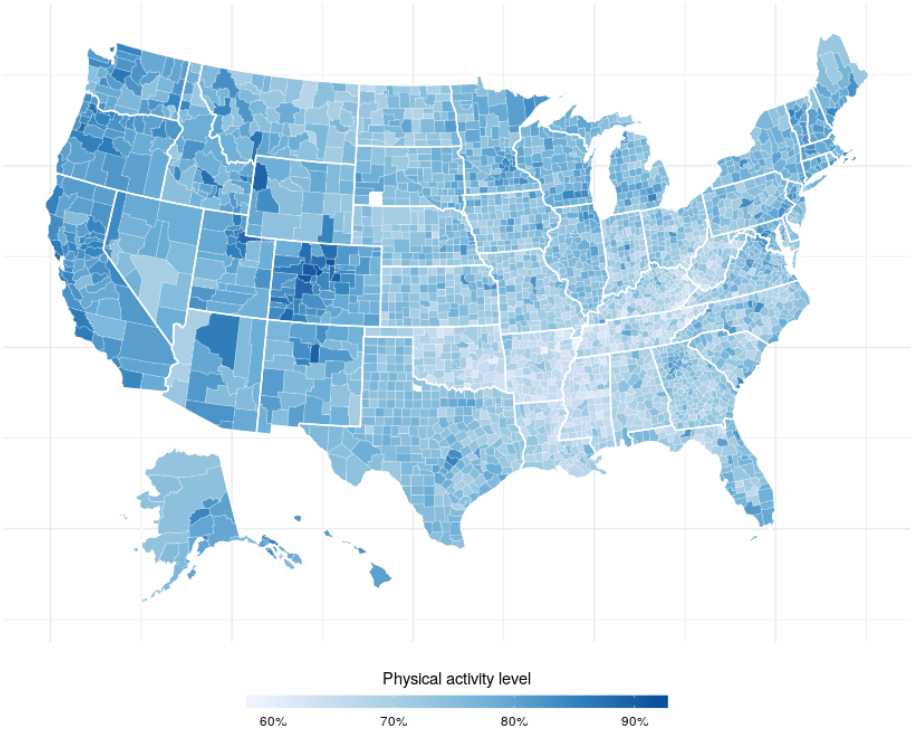


Figure 2. Map of physical activity–related geolocated tweets in the United States.



XSL•FO
RenderX

Figure 3. Map of Gini index across the United States.

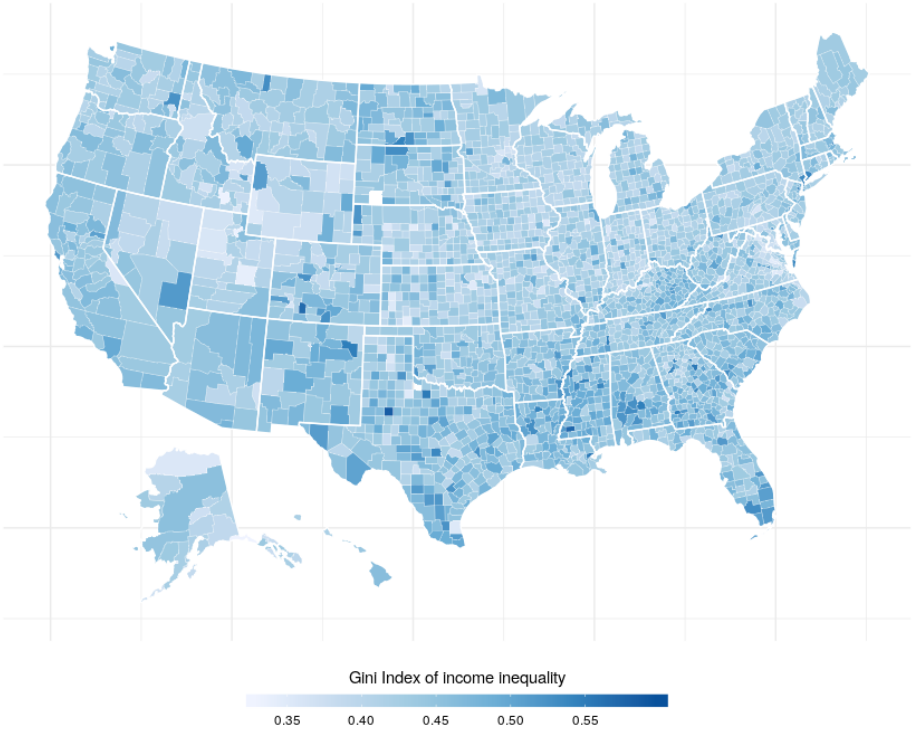
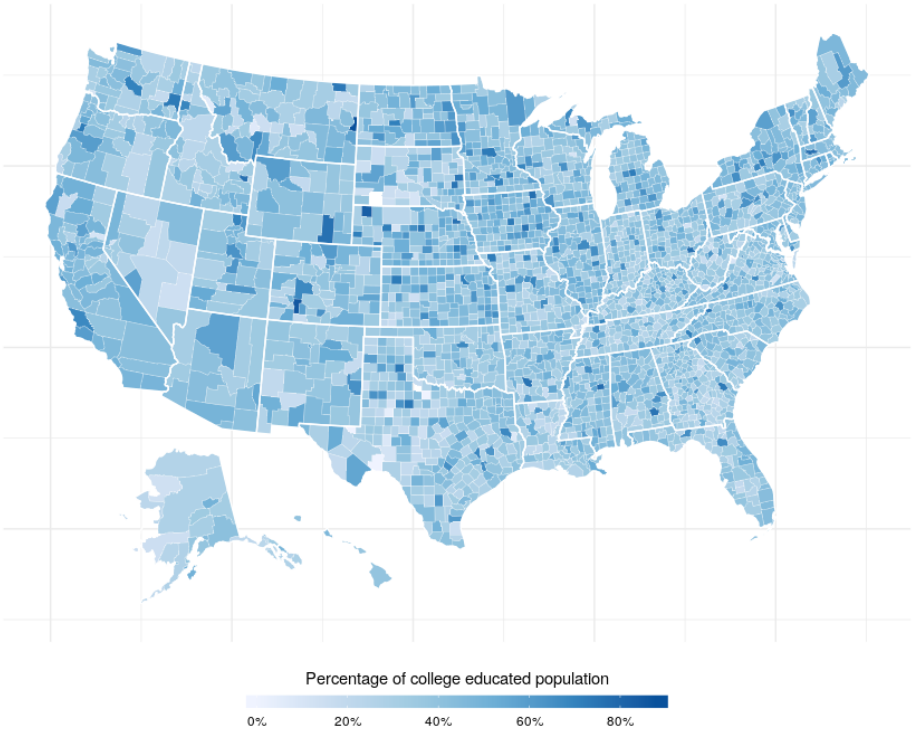


Figure 4. Map of education level across the United States.



XSL•FO
RenderX

Textbox 1. Example tweets classified with positive, negative, and neutral sentiment.

Classified positive:

- "Growth physically, mentally, spiritually, financially. That will be my 2017....#instafit…"
- "I think between us we've lost a whole person! #fitgoals…"
- "Today's reminder: it's about practice, not perfection! #yogajournal…"

Classified negative:

- "I hate Tuesdays. Extra. #cardio"
- "Shoutout to #crossfit ... More like curbstomp am I right guys?! @ Portland, Oregon"
- "Back to the grind… #cardio"

Classified neutral:

- "Where's the #belly I see it! #handful but in a good way #fitness #goals continue…"
- "Day 2 of my 2 a day workouts down #gymlife #planetfitness…"
- "Clips of my leg day #squatsanddeads #bodybuilding #powerlifting #strongman #olympiclicting #fit…"

Table 1. Summary of county-level physical activity, activity-related tweets, and the Gini index.

Variable	Physically active, %	Physically active, % (age adjusted)	Gini index	Education	Physical activity tweets	The ratio between positive and negative physical activity tweets
Physically active, %	1	0.99[a]	–0.16[a]	0.26[a]	0.38[a]	0.13[a]
Physically active, % (age adjusted)	—[b]	1	–1.77[a]	0.24[a]	0.34[a]	0.10[a]
Gini index	—	—	1	0.04[c]	0.05[c]	0.09[a]
Education	—	—	—	1	0.22[a]	0.16[a]
Physical activity tweets	—	—	—	—	1	0.20[a]
Positive / negative physical activity tweets ratio	—	—	—	—	—	1

[a]$P<.001$.

[b]Not applicable.

[c]$P<.02$.

The regression models showed that the percentage of physical activity–related tweets was significantly associated with the physical activity level (Table 2) and age-adjusted physical activity on a county level (Table 3) while adjusting for both the Gini index and education level. However, the sentiment of the physical activity–related tweets was not a significant predictor of the physical activity level and age-adjusted physical activity on a county level after including the Gini index and education level in the model. The best-fit model for predicting county-level physical activity incorporated the percentage of physical activity–related tweets, the Gini index, and the prevalence of college education. However, the overall explained variance of the model was low (R^2=.11) Similarly, the best-fit model for predicting county-level physical activity (R^2=.09) after adjusting for age used the percentage of physical activity–related tweets, the Gini index, and the education level.

XSL•FO
RenderX

Table 2. Regression analysis for physical activity–related tweets and county-level physical activity level.

Variables	beta	SE	*P* value
Model 1[a]			
Gini index	−0.16	2.54	<.001
Education	0.26	.01	<.001
Percent of physical activity–related tweets	0.11	.20	<.001
Model 2[b]			
Gini index	−0.12	3.78	<.001
Education	2.95	.01	<.001
Sentiment of physical activity–related tweets (positive/negative ratio)	−0.01	.37	.56
Model 3[c]			
Gini index	−0.12	3.79	<.001
Education	0.30	0.01	<.001
Percent of physical activity–related tweets	0.05	.37	.02
Sentiment of physical activity–related tweets (positive/negative ratio)	−0.01	.24	.53

[a]$F_{3,3137}$=116.30; P<.001; R^2=.11.

[b]$F_{3,1704}$=55.99; P<.001; R^2=.09.

[c]$F_{3,1704}$=43.517; P<.001; R^2=.09.

Table 3. Regression analysis for physical activity–related tweets and age-adjusted county-level physical activity level.

Variables	beta	SE	*P* value
Model 1[a]			
Gini index	−0.18	2.44	<.001
Education	0.23	.01	<.001
Percent of physical activity–related tweets	0.10	0.20	<.001
Model 2[b]			
Gini index	−0.13	3.63	<.001
Education	0.25	.01	<.001
Sentiment of physical activity–related tweets (positive/negative ratio)	−0.02	.35	.44
Model 3[c]			
Gini index	−0.13	3.64	<.001
Education	0.26	.01	<.001
Percent of physical activity–related tweets	0.05	.23	.03
Sentiment of physical activity–related tweets (positive/negative ratio)	−0.02	.35	.41

[a]$F_{3,3137}$=102.93; P<.001; R^2=.10.

[b]$F_{3,1704}$=43.09; P<.001; R^2=.07.

[c]$F_{3,1704}$=33.52; P<.001; R^2=.07.

Discussion

Principal Findings

This study evaluated the feasibility of using Twitter data to monitor physical activity levels by assessing whether geotagged conversations about physical activity behaviors can be extracted from Twitter and whether physical activity–related tweets could be used to monitor physical activity levels. Results suggest that it is feasible to extract physical activity–related geotagged conversations from Twitter. Furthermore, the results suggest that there was a significant association between physical activity–related tweets and physical activity levels while accounting for the Gini index of income inequality, population, and education on a county level across the United States. However, the overall association between physical

XSL•FO
RenderX

activity–related tweets and physical activity levels on a county level was weak.

Research Implications

Exploring the relationship between physical activity–related tweets and physical activity levels on a county level has several important research implications. First, these findings support the continued research in using nontraditional data sources, such as social media data, to monitor physical activity–related behaviors. Second, our results demonstrated a potential application for using social media data as a complementary tool to aid in both historical and real-time tracking of population-level physical activity. A strength of this study is controlling for related demographic factors such as income inequality and education in various geographic locations in our model. Finally, physical activity researchers can build upon the methods used in this study to find new methods of using social media data to monitor physical activity outcomes. Physical activity researchers can leverage these social media analysis techniques to build models that can predict physical activity levels in real-time. The analysis methods used in this study could in the future aid public health agencies in identifying particular physical activity–related trends or geographical areas of concern on which to focus their health and wellness initiatives.

Findings from this study validate and extend previously published work that the content of the tweets can be potentially used to monitor and predict behavior and health outcomes [10,16,22]. It is worth noting that even though we did not show a significant association between the sentiment of physical activity–related tweets and age-adjusted physical activity on a county level, previous studies have shown that the sentiment of the tweets can be used to predict health outcomes. Specifically, a previous study reported that positive sentiment tweets were moderately correlated with lower obesity rates in 190 US MSAs [28]. These findings suggest that sentiment analysis may not be an appropriate estimator of physical activity level on a county level but may still be an appropriate estimator in other health-related outcomes on an MSA level.

Although infoveillance or infodemiology studies such as this are important to epidemiology to avoid *ecological fallacies* [19], it is critical for future research to examine the relationship between social media data and physical activity level on an individual level. Studying the prevalence of physical activity is a complex and nuanced topic, one that may be strongly influenced by an individual's surrounding environment. We were able to obtain improved model performance through the inclusion of per county Gini index data on income inequality and education level. However, future studies will need to investigate whether other known metrics or indicators (eg, the built environment) can be incorporated into Twitter data to create models with improved accuracy in predicting physical activity.

Limitations

There are several limitations in the study. There was a lag time and time frame disparity between the Twitter data and physical activity data. The most recent county-level physical activity data, collected as part of its annual BRFSS surveys, was from 2014. In addition, there exist inherent biases that must be noted for any sampling of geotagged Twitter data. Studies on demographics on the platform have found a skew toward a younger, wealthier demographic in general [21], as well as increased representation from minority groups and urban populations when looking at geotagged tweets in particular [30]. This means that the observed relationship between physical activity–related tweets and physical activity needs to be interpreted with caution as certain demographic or regional groups may be predisposed toward a certain physical activity level. Nevertheless, this study was a feasibility study primarily designed to evaluate whether social media conversations that suggest physical activity–related behaviors could be extracted and used to monitor physical activity at a population level. Second, we only used one source of social media data (Twitter), thus limiting the generalizability of our findings. Using Twitter API's *garden hose* approach with geographic filtering also limits data collection to less than a 1% random sample of all tweets posted in any given time frame. Future studies will need to explore whether model accuracy can be improved using multiple data sources (eg, Twitter, Facebook, and Instagram) to exploit user overlap between certain social media platforms as well as over a longer data collection period. Finally, studies of this nature rely heavily on the accuracy of the classifier for labeling physical activity–related tweets and conducting sentiment analysis. In particular, the finer-grained filtering offered by sentiment analysis did not appear to offer a notable improvement in fit or classification accuracy. It should be noted that this failure of complex or synthesized features to improve model quality has been observed earlier [28]. Subsequent studies may supplant our list-of-hashtags classifier with machine-learning classification approaches to potentially discover keywords, text structure, or other features that may be used to boost both precision and recall, as well as attempt using state-of-the-art sentiment analysis techniques to construct and train custom classifiers that are a better fit for this specific subset of Twitter data.

Moving forward, there are still other possible features to extract from Twitter data that may be tested for association with levels of physical activity. Although this study focused exclusively on filtering Twitter data by keywords and conducting sentiment analysis, there may be other natural language processing techniques that could be applied to the dataset [16,36]. Future research could investigate training predictive models on a larger, longitudinal dataset of both tweets and physical activity data. If successful, such models could be leveraged to effectively predict levels of physical activity and inactivity using social media data.

Conclusions

This study evaluated the feasibility of using social media data to monitor physical activity levels on a county-by-county basis. Results from this study suggest that it is feasible to identify geotagged physical activity–related conversations from Twitter data and link them to population-based physical activity outcome data for analyses. We found that the conversation from tweets was weakly associated with county-level physical activity levels in the United States. Future research can build on the methods

XSL•FO
RenderX

used in this study to further refine the models that use real-time social media data to monitor physical activity levels.

Acknowledgments

This work is supported by the Internal Research Creative Project Grant at the University of Victoria.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Hashtags used to identify tweets related to physical activity.

[DOCX File, 14KB - jmir_v21i6e12394_app1.docx]

References

1. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
2. Liu S, Goodman J, Nolan R, Lacombe S, Thomas SG. Blood pressure responses to acute and chronic exercise are related in prehypertension. Med Sci Sports Exerc 2012 Sep;44(9):1644-1652. [doi: 10.1249/MSS.0b013e31825408fb] [Medline: 22899388]
3. Lacombe SP, Goodman JM, Spragg CM, Liu S, Thomas SG. Interval and continuous exercise elicit equivalent postexercise hypotension in prehypertensive men, despite differences in regulation. Appl Physiol Nutr Metab 2011 Dec;36(6):881-891. [doi: 10.1139/h11-113] [Medline: 22070641]
4. Umpierre D, Ribeiro P, Kramer C, Leitão CB, Zucatti A, Azevedo M, et al. Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis. J Am Med Assoc 2011 May 4;305(17):1790-1799. [doi: 10.1001/jama.2011.576] [Medline: 21540423]
5. Friedenreich C, Orenstein M. Physical activity and cancer prevention: etiologic evidence and biological mechanisms. J Nutr 2002 Dec;132(11 Suppl):3456S-3454S. [doi: 10.1093/jn/132.11.3456S] [Medline: 12421870]
6. PAGA. Physical Activity Guidelines Advisory Committee Report. Washington, DC: US Department of Health and Human Services; 2008.
7. Centers for Disease Control and Prevention. Facts about Physical Activity URL: https://www.cdc.gov/physicalactivity/data/facts.htm [accessed 2019-05-15] [WebCite Cache ID 78ONZdpa5]
8. Caspersen CJ, Pereira MA, Curran KM. Changes in physical activity patterns in the United States, by sex and cross-sectional age. Med Sci Sports Exerc 2000 Sep;32(9):1601-1609. [Medline: 10994912]
9. Trost SG, Owen N, Bauman AE, Sallis JF, Brown W. Correlates of adults' participation in physical activity: review and update. Med Sci Sports Exerc 2002 Dec;34(12):1996-2001. [doi: 10.1249/01.MSS.0000038974.76900.92] [Medline: 12471307]
10. Liu S, Young S. A survey of social media data analysis for physical activity surveillance. J Forensic Leg Med 2018 Jul;57:33-36. [doi: 10.1016/j.jflm.2016.10.019] [Medline: 29801949]
11. Liu S, Dunford SD, Leung YW, Brooks D, Thomas SG, Eysenbach G, et al. Reducing blood pressure with internet-based interventions: a meta-analysis. Can J Cardiol 2013 May;29(5):613-621. [doi: 10.1016/j.cjca.2013.02.007] [Medline: 23618507]
12. Liu S, Brooks D, Thomas SG, Eysenbach G, Nolan RP. Effectiveness of user-and expert-driven web-based hypertension programs: an RCT. Am J Prev Med 2018 Dec;54(4):576-583. [doi: 10.1016/j.amepre.2018.01.009] [Medline: 29456025]
13. Liu S, Hodgson C, Zbib AM, Payne AY, Nolan RP. The effectiveness of loyalty rewards to promote the use of an Internet-based heart health program. J Med Internet Res 2014 Jul;16(7):e163 [FREE Full text] [doi: 10.2196/jmir.3458] [Medline: 24989982]
14. Nolan RP, Liu S, Shoemaker JK, Hachinski V, Lynn H, Mikulis DJ, et al. Therapeutic benefit of internet-based lifestyle counselling for hypertension. Can J Cardiol 2012 May;28(3):390-396. [doi: 10.1016/j.cjca.2012.02.012] [Medline: 22498181]
15. Valle C, Tate D, Mayer D, Allicock M, Cai J. A randomized trial of a Facebook-based physical activity intervention for young adult cancer survivors. J Cancer Surviv 2013 Sep;7(3):355-368 [FREE Full text] [doi: 10.1007/s11764-013-0279-5] [Medline: 23532799]
16. Salathé M. Digital epidemiology: what is it, and where is it going? Life Sci Soc Policy 2018 Jan 4;14(1):1 [FREE Full text] [doi: 10.1186/s40504-017-0065-7] [Medline: 29302758]
17. Perrin A. Pew Research Center. 2015. Social Media Usage URL: https://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/ [accessed 2019-05-15] [WebCite Cache ID 78ONsKmKK]

XSL•FO
RenderX

18. Young SD. Recommended guidelines on using social networking technologies for HIV prevention research. AIDS Behav 2012 Oct;16(7):1743-1745 [FREE Full text] [doi: 10.1007/s10461-012-0251-9] [Medline: 22821067]
19. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
20. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
21. Greenwood S, Perrin A, Duggan M. Pew Research Center. 2017. Social Media Update 2016 URL: https://www.pewinternet.org/2016/11/11/social-media-update-2016/ [accessed 2019-05-15] [WebCite Cache ID 78OO53jLF]
22. Liu S, Zhu M, Yu DJ, Rasin A, Young SD. Using real-time social media technologies to monitor levels of perceived stress and emotional state in college students: a web-based questionnaire study. JMIR Ment Health 2017 Jan 10;4(1):e2 [FREE Full text] [doi: 10.2196/mental.5626] [Medline: 28073737]
23. Liu S, Zhu M, Young SD. Monitoring freshman college experience through content analysis of tweets: observational study. JMIR Public Health Surveill 2018 Jan 11;4(1):e5 [FREE Full text] [doi: 10.2196/publichealth.7444] [Medline: 29326096]
24. Young SD, Rivers C, Lewis B. Methods of using real-time social media technologies for detection and remote monitoring of HIV outcomes. Prev Med 2014 Jun;63:112-115 [FREE Full text] [doi: 10.1016/j.ypmed.2014.01.024] [Medline: 24513169]
25. Bodnar TS. Validating models for disease detection using twitter. 2013 May Presented at: WWW'13 Companion; May 13-17, 2013; Rio de Janeiro, Brazil p. 699-702 URL: http://dl.acm.org/citation.cfm?id=2487788.2488027
26. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. PLoS One 2013;8(12):e83672 [FREE Full text] [doi: 10.1371/journal.pone.0083672] [Medline: 24349542]
27. 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]
28. Gore RJ, Diallo S, Padilla J. You are what you tweet: connecting the geographic variation in america's obesity rate to Twitter content. PLoS One 2015;10(9):e0133505 [FREE Full text] [doi: 10.1371/journal.pone.0133505] [Medline: 26332588]
29. Nguyen Q, Li D, Meng H, Kath S, Nsoesie E, Li F, et al. Building a national neighborhood dataset from geotagged Twitter data for indicators of happiness, diet, and physical activity. JMIR Public Health Surveill 2016 Oct 17;2(2):e158 [FREE Full text] [doi: 10.2196/publichealth.5869] [Medline: 27751984]
30. Paul M, Dredze M. You are what you tweet: Analyzing twitter for public health. 2011 Presented at: Fifth International AAAI Conference on Weblogs and Social Media; July 17–21, 2011; Barcelona, Catalonia, ES.
31. Manley A. Physical Activity And Health: A Report Of The Surgeon General. Darby, PA: Diane Pub Co; 2019.
32. Baziotis C, Pelekis N, Doulkeridis C. DataStories at SemEval-2017 Task 4: Deep LSTM with Attention for Message-level and Topic-based Sentiment Analysis. 2017 Presented at: International Workshop on Semantic Evaluation 2011; August 3-4, 2017; Vancouver, Canada p. 747.
33. Bethard S, Carpuat M, Cer D, Jurgens D, Nakov P, Zesch T. Association for Computational Linguistics. 2016. Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016) URL: https://www.aclweb.org/anthology/S16-1000 [accessed 2019-05-17] [WebCite Cache ID 78QcxPpyr]
34. Rocktäschel T, Grefenstette E, Hermann K, Kočiský T, Blunsom P. arXiv. 2015. Reasoning about entailment with neural attention URL: https://arxiv.org/abs/1509.06664 [accessed 2019-05-17] [WebCite Cache ID 78Qd71Did]
35. Giles-Corti B, Donovan RJ. Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment. Prev Med 2002 Dec;35(6):601-611. [Medline: 12460528]
36. Kumar S. Twitter Data Analytics. New York: Springer; 2014.

Abbreviations

ACSM: American College of Sports Medicine
API: application programming interface
BRFSS: Behavioral Risk Factor Surveillance System
MSA: Metropolitan Statistical Area
SemEval: Semantic Evaluation

XSL•FO
RenderX

Edited by G Eysenbach; submitted 16.10.18; peer-reviewed by N Stetten, R Gore; comments to author 03.02.19; revised version received 12.04.19; accepted 14.04.19; published 03.06.19.

Please cite as:
Liu S, Chen B, Kuo A
Monitoring Physical Activity Levels Using Twitter Data: Infodemiology Study
J Med Internet Res 2019;21(6):e12394
URL: https://www.jmir.org/2019/6/e12394/
doi:10.2196/12394
PMID:31162126

©Sam Liu, Brian Chen, Alex Kuo. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Sentiment Analysis of Social Media on Childhood Vaccination: Development of an Ontology

Jeongah On[1], RN; Hyeoun-Ae Park[1,2], RN, PhD, FAAN; Tae-Min Song[3], PhD

[1]College of Nursing, Seoul National University, Seoul, Republic of Korea

[2]Research Institute of Nursing Science, Seoul National University, Seoul, Republic of Korea

[3]Department of Health Management, Sahmyook University, Seoul, Republic of Korea

Corresponding Author:
Hyeoun-Ae Park, RN, PhD, FAAN
College of Nursing, Seoul National University
103, Daehak-ro, Jongno-gu
Seoul, 03080
Republic of Korea
Phone: 82 27408827
Fax: 82 27654103
Email: hapark@snu.ac.kr

Abstract

Background: Although vaccination rates are above the threshold for herd immunity in South Korea, a growing number of parents have expressed concerns about the safety of vaccines. It is important to understand these concerns so that we can maintain high vaccination rates.

Objective: The aim of this study was to develop a childhood vaccination ontology to serve as a framework for collecting and analyzing social data on childhood vaccination and to use this ontology for identifying concerns about and sentiments toward childhood vaccination from social data.

Methods: The domain and scope of the ontology were determined by developing competency questions. We checked if existing ontologies and conceptual frameworks related to vaccination can be reused for the childhood vaccination ontology. Terms were collected from clinical practice guidelines, research papers, and posts on social media platforms. Class concepts were extracted from these terms. A class hierarchy was developed using a top-down approach. The ontology was evaluated in terms of description logics, face and content validity, and coverage. In total, 40,359 Korean posts on childhood vaccination were collected from 27 social media channels between January and December 2015. Vaccination issues were identified and classified using the second-level class concepts of the ontology. The sentiments were classified in 3 ways: positive, negative or neutral. Posts were analyzed using frequency, trend, logistic regression, and association rules.

Results: Our childhood vaccination ontology comprised 9 superclasses with 137 subclasses and 431 synonyms for class, attribute, and value concepts. *Parent's health belief* appeared in 53.21% (15,709/29,521) of posts and positive sentiments appeared in 64.08% (17,454/27,236) of posts. Trends in sentiments toward vaccination were affected by news about vaccinations. Posts with *parents' health belief*, *vaccination availability*, and *vaccination policy* were associated with positive sentiments, whereas posts with *experience of vaccine adverse events* were associated with negative sentiments.

Conclusions: The childhood vaccination ontology developed in this study was useful for collecting and analyzing social data on childhood vaccination. We expect that practitioners and researchers in the field of childhood vaccination could use our ontology to identify concerns about and sentiments toward childhood vaccination from social data.

(J Med Internet Res 2019;21(6):e13456) doi:10.2196/13456

KEYWORDS

social media; vaccination; health information interoperability; semantics

XSL•FO
RenderX

Introduction

Vaccination is an effective means of inducing active immunity against infection by administering a vaccine made by killing or weakening the pathogenicity of microorganisms. In addition to preventing individuals from becoming ill, this creates herd immunity, thus preventing transmission of infection through social groups [1]. Childhood vaccination, starting at infancy when the immune system is weak, is very important for immunogenesis and disease prevention. Therefore, the government of each country sets out schedules for childhood vaccinations and recommends vaccination at the appropriate time. The Korean Centers for Disease Control and Prevention (KCDC) recommends that parents vaccinate their children aged younger than 12 years according to a standard vaccination schedule and supports vaccination with policies, such as the provision of financial support [2].

According to a national childhood vaccination coverage survey conducted by KCDC in 2016, the vaccination rate of children aged 3 years was 96.5% [3], which is higher than the herd immunity threshold. However, there are concerns raised by the public that vaccination might not be effective in preventing infectious disease, especially when there were intermittent outbreaks of vaccine-preventable diseases, such as viral hepatitis A and varicella in South Korea. The public questions about the safety of vaccines because of bogus rumors, such as the Measles, Mumps, and Rubella (MMR) vaccines causing autism and thimerosal, a mercury-based preservative contained in some vaccines, causing brain damage [4].

The number of posts on social media claiming that vaccination is not necessary or even harmful is increasing. These negative sentiments toward vaccination may affect people's intention to vaccinate and thus lead to a reduction in vaccination rates [5-9]. Therefore, to maintain vaccination rates above the herd immunity threshold, it is necessary to monitor public concerns about and sentiments toward vaccination and identify factors affecting them.

Several studies have assessed public knowledge and perceptions of vaccination, intent to vaccinate, and factors affecting vaccination intent or behavior [5-8,10-12]. Most of these studies were conducted using mail or telephone surveys or personal interviews. However, there are disadvantages of these survey studies, such as long research time, small sample size, representativeness of sample, low response rates, and interviewer bias.

Social networking services (SNSs) are emerging as a medium that can be used to identify public concerns and sentiments in various fields [13]. Social data are used to identify concerns about vaccination, such as their safety and side effects [14-16], as well as sentiments toward vaccines in general [14] or toward specific vaccines, such as the MMR vaccine [17], influenza A vaccination [18], and Human Papilloma Virus (HPV) vaccine [19]. In addition, social data are used to identify correlations between sentiments toward vaccines and epidemics of vaccine-preventable diseases [17,18] and vaccine information flow [18]. In South Korea, no research has yet examined concerns about and sentiments toward vaccines and vaccination using social data.

The internet usage rate of Koreans is as high as 90.3% and 68.2% of them use social media [20]. Specifically, the usage rate of those aged 20 to 30 years who can be parents of young children, is as high as 89.5% [20]. In fact, there are about 45,000 parenting communities in the online cafes of the 2 most popular Web portals in South Korea, Daum and Naver. Thus, data on childhood vaccination can be obtained, and vaccination issues and sentiments can be identified. Vaccination issues in this study were defined as opinions, perceptions, concerns, and worries about vaccines and vaccination posted on social media. Sentiments in this study were defined as feeling and emotion on childhood vaccination posted on social media.

Social media posts are unstructured data composed mostly of text. To use such unstructured data for analysis, a hierarchical classification of terms, relationships of terms, such as synonyms and hyponyms, and clustering based on the frequency of terms are used [21,22]. However, these methods are not sufficient for understanding the semantics of terms [21,22]. An ontology defining the meanings and inherent attributes of concepts, capturing relationships between them, and containing terms covering thesaurus, is required for social data analysis to solve this issue [21,23-25]. An ontology can help researchers understand the semantics of and the relationships between concepts when contextual knowledge is lacking. Terms included in the ontology help researchers collect the social data that appear in the form of various synonyms. Unfortunately, no ontology is yet available that can be used to identify vaccination issues and sentiments toward childhood vaccination.

Therefore, we developed an ontology for use as a framework for collecting and analyzing social data on childhood vaccination. We used this ontology to identify vaccination issues and sentiments toward childhood vaccination, trends in these sentiments, and relationships between vaccination issues and sentiments in social data posted in Korean.

Methods

Development of a Childhood Vaccination Ontology

We developed a childhood vaccination ontology by following the Ontology Development 101 [26]. Ontology development consists of the following 5 steps, and it was an iterative process.

Step 1: Determining the Domain and Scope of the Ontology

To determine the domain and scope of the ontology, we first created competency questions [27], which is a list of questions that the childhood vaccination ontology should be able to answer, for example: *What are the childhood vaccination items?* The competency questions were also used for ontology evaluation in the final step.

Step 2: Consideration of Reuse of Existing Ontologies

We searched for existing ontologies and conceptual frameworks related to vaccination in BioPortal [28], a repository of biomedical ontologies and research papers. We identified the Vaccine Ontology [29], which contains classifications of

XSL•FO
RenderX

vaccines and vaccine components, vaccine quality and phenotypes, and host immune responses to vaccines. We also identified 5 models of vaccination decisions and hesitancy [30-34]. They are as follows: (1) a conceptual model of the role of parental attitudes and beliefs in decision making about child and adolescent vaccination [30], (2) a conceptual framework for HPV vaccine acceptance and adherence, focusing on sociocultural factors impacting vaccine adherence behavior [31], (3) a framework for determinants of H1N1 influenza vaccine uptake utilizing the social ecological model [32], (4) a model for assessing determinants of vaccine hesitancy in different settings [33], and (5) a conceptual model of determinants of individual decision making about vaccination [34]. We reviewed whether the Vaccine Ontology and the models of vaccination decisions and hesitancy could be used to develop a childhood vaccination ontology.

Step 3: Collecting Terms and Extracting Concepts

We collected terms within the domain and scope of the ontology by reviewing 3 vaccination practice guidelines developed by the KCDC [35], the United States Centers for Disease Control and Prevention [36], and the Public Health England [37], and 101 research papers found by searching for the keywords *vaccination* and *immunization*. In addition, we searched SNSs to identify new terms that were not collected from practice guidelines or research papers. We extracted class concepts from the terms.

Step 4: Developing the Ontology and Terminology

The class hierarchy was developed using a top-down approach. The superclasses of the ontology and their relationships were constructed by integrating 2 models of vaccination decisions and hesitancy [31,32]. Subclasses were defined by grouping of the vaccination-related concepts from the 5 models of vaccination decisions and hesitancy [30-34]. The extracted concepts were arranged by mapping or adding to the class hierarchy. We represented class concepts as entity-attribute-value (EAV) models. We also developed a terminology, including synonyms for class, attribute, and attribute value concepts. The ontology was developed using the Protégé 5.0.0 ontology editor.

Step 5: Evaluating the Ontology

We evaluated the ontology with description logic (DL) verification, face and content validity, and coverage evaluation. DL of ontology was verified by whether the ontology provided appropriate answers to the competency questions developed earlier. We converted competency questions into Protégé DL queries (DL-query) using the class concepts and relationships between classes and tested whether it generated the correct answer upon entering a DL-query. For example, the competency question, *What are the adverse events experienced after vaccination?* was converted to the DL-query, *hasType some experience of vaccine adverse event*. After entering this query into Protégé, we tested whether the answers to the query, such as *edema* and *hypersensitivity*, were adverse events associated with the vaccine.

Face validity of the ontology was evaluated by 3 health informatics experts in biomedical ontology. They were asked to assess superficially and subjectively assess whether the ontology is valid for identifying childhood vaccination issues. Content validity of ontology was evaluated by 3 domain experts with Master's degree and more than 4 years of experience in pediatric nursing. They were asked to rate the ontology classes on a 4-point scale (1=very invalid, 2=invalid, 3=valid, and 4=very valid) as to whether it is valid for identifying childhood vaccination issues. The content validity index (CVI) of the ontology was calculated by taking the average of the class-level CVIs, which were computed by dividing the number of experts with 3 to 4 points by the total number of experts for each class.

Coverage of the ontology was examined by comparing terms extracted from social data with the concepts and synonyms of the ontology. We revised the ontology by adding new terms.

Analyzing Social Data on Childhood Vaccination

We collected posts on childhood vaccination posted in the Korean language from 27 social media channels between January 1 and December 31, 2015 (see Multimedia Appendix 1). The channels included 1 SNS (Twitter, a popular microblogging site), 2 online cafes (Daum Café and NAVER Café, 2 popular online community services), 4 internet blogs (eg, the NAVER blog), and 20 message boards (eg, NATE Pann). We used *vaccination* as a search keyword and *vaccine injection*, *child vaccination*, *child vaccine injection*, *infant vaccination*, *infant vaccine injection*, *toddler vaccination*, and *toddler vaccine injection* as synonyms. Posts on livestock or plant vaccination were excluded. Data were collected by Smart Insight [38], a big-data marketing platform. The social data do not have any identifiable personal information, such as user profiles. This study was approved by the Institutional Review Board of the Korea Institute for Health and Social Affairs.

We preprocessed the collected social data by treating the analysis unit as a single post. After extracting the terms from each post, we identified the terms related to vaccination as vaccination issues and the emotional words as sentiments. The terms identified as vaccination issues were classified into the second-level classes of the ontology. The emotional words identified as sentiments were classified into positive and negative emotional words using a *universal emotion keyword list* developed by Smart Insight. We counted the numbers of positive and negative emotional words in each post. When the number of positive emotional words was more than the number of negative emotional words, the post was classified as a positive sentiment. When the number of negative emotional words was more than the number of positive emotional words, the post was classified as a negative sentiment. When the number of positive emotional words was equal to the number of negative emotional words, the post was classified as a neutral sentiment.

We analyzed the frequencies of vaccination issues and sentiments. We compared monthly sentiment trends with selected vaccination-related news. Vaccination-related news was selected from news articles on vaccines, vaccination, and infectious diseases, searched from Naver, which is one of the largest news portals in South Korea. We performed sentiment analysis by conducting logistic regression analysis and association analysis. We used logistic regression analysis to investigate how vaccination issues affect sentiments. Significant

XSL•FO
RenderX

vaccination issues from univariate analyses ($P<.05$) were used as independent variables and sentiments as the dependent variable, with positive and negative sentiments converted to *1* and *0*, respectively. We used association analysis applying a priori principle algorithm to investigate which sets of vaccination issues were associated with sentiments. The interestingness of the association rules are expressed as support, confidence, and lift. Support means proportion of posts with set of vaccination issues and sentiments in the entire posts. Confidence means proportion of posts with sentiments among posts with set of vaccination issues. Lift means ratio of probability of sentiments when set of vaccination issues is appeared to probability of sentiments when set of vaccination issues is not appeared. We restricted the consequences of rules to positive and negative sentiments. A frequent item set was defined as rules satisfying minimal support (0.05) and confidence (0.5) constraints. We used the R software package (version 3.3.1) for both logistic regression and association analyses.

Results

Development of a Childhood Vaccination Ontology

We determined the domain and scope of the ontology from personal perceptions, behavior, and experiences of vaccination as well as social, environmental, institutional, and political factors related to vaccination based on 21 competency questions (see Textbox 1) created. We restricted vaccinations to those recommended by the KCDC for children aged 0 to 12 years.

A total of 5 models of vaccination decisions and hesitancy [30-34] were used in the childhood vaccination ontology development. We modified the domain and scope of the ontology by renaming personal perceptions as parent factor, and adding child factor, family and friend factor, organizational factor, and community factor related to vaccination.

We collected 883 terms covering the domain and scope of ontology and extracted 133 unique class concepts from these terms. We defined hierarchical and attribute relationships between classes based on 5 models of vaccination decisions and hesitancy [30-34]. Various factors affect vaccine uptake, decision making, and hesitancy [30,32-34]. Vaccination experience is also an important factor affecting completion of the multistep vaccination schedule [31]. We viewed vaccination as a process progressing from *before vaccination* and *vaccination* onto *after vaccination*. Based on the Social Ecological Framework for H1N1 Influenza Vaccine Uptake in the United States [32], we placed *child*, *parent*, *family and friend*, *organization and institution*, *society and community*, and *policy* levels of factors affecting vaccination before vaccination. We also placed *vaccination intention* before vaccination. *Vaccination behavior* was placed at vaccination and *vaccination experience* was placed after vaccination. Various levels of factors affect vaccination intention, behavior, and experience. Vaccination intention affects vaccination behavior, and vaccination behavior affects vaccination experience. Vaccination experience, in turn, affects vaccination intention for the next round of vaccinations.

Textbox 1. A list of competency questions.

What are the childhood vaccination items?
Where is the child vaccinated?
How many vaccinations does it take?
Which body sites are vaccinated?
When should vaccination be appropriate?
What is the cost of vaccination?
Do parents know about vaccination?
Where do parents get vaccination information?
What are the vaccination policies?
What influences vaccination decisions?
What do parents think about the child's side when deciding on vaccination?
Which parents vaccinate their child?
What are the parents' health belief that influence vaccination decisions?
What previous vaccination experience influenced vaccination decisions?
What are the family and friend side factors that influence vaccination decisions?
What agencies are involved in vaccination or vaccine?
What are the epidemics affecting vaccination decisions?
What are the media reports that influence vaccination decisions?
What are the adverse events experienced after vaccination?
What is the experience of medical staff involved in vaccination?
What are the sentiments about vaccination?

XSL•FO
RenderX

Figure 1. Childhood vaccination ontology up to the second-level class.

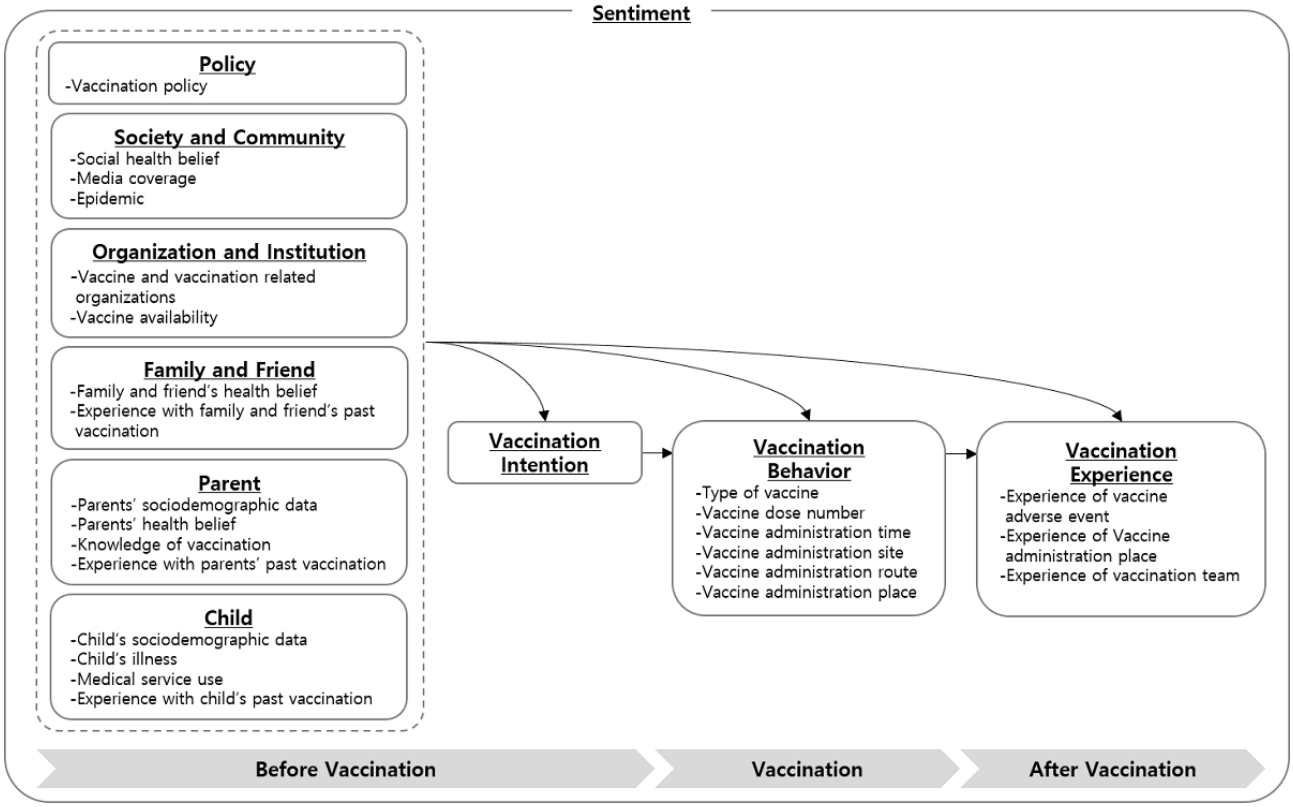


Figure 1 shows the childhood vaccination ontology (up to the second level) that we developed. The ontology consists of 9 superclasses: *child, parent, family and friend, organization and institution, society and community, policy, vaccination intention, vaccination behavior,* and *vaccination experience*. *The child* class included child-related factors for vaccination, such as age, illness, and experience with past vaccination. *The parent* class included parent-related factors for vaccination, such as health belief and knowledge about vaccination. *The family and friend* class included family- and friend-related factors for vaccination, such as family and friend's health belief. *The organization and institution* class included vaccine- and vaccination-related organizations and vaccine availability. *The society and community* class included media coverage and epidemics. *The policy* class included vaccination policy. *The vaccination intention* class included decisions on whether to vaccinate in future. *The vaccination behavior* class included the types and vaccine dose number of the vaccination administered. *The vaccination experience* class included experience after vaccination, such as adverse events. The depths of the superclasses varied from 1 level for *vaccination intention* to 4 levels for *parent*.

We developed EAV models of the 103 lowest level class concepts. For example, *experience of vaccine adverse event* had attributes of *hasType* and *hasSeverity*. Of these, *hasType* had values, such as *pain* and *edema*, and *hasSeverity* had values, such as *mild*, *moderate*, and *severe*. We also developed a terminology composed of 126 synonyms for 133 classes, 1 synonym for 12 attributes, and 343 synonyms for 268 values.

With the DL verification, it was found that the ontology correctly answered all 21 competency questions. With the face validity, it was found that the ontology was superficially and subjectively valid for identifying childhood vaccination issues. With the content validity, it was found that the CVI of the ontology was 1.0 and all classes of the ontology were valid for identifying childhood vaccination issues. With coverage evaluation, 138 of the 148 (93.2%) concepts were covered by the ontology and 163 out of 575 (28.3%) synonyms were covered by the ontology. We revised the ontology by adding 10 value concepts, 288 synonyms for classes, and 124 synonyms for values. Finally, the childhood vaccination ontology was composed of 133 classes, 414 synonyms for 133 classes, 1 synonym for 12 attributes, and 467 synonyms for 278 values.

Analyzing Social Data on Childhood Vaccination

Frequency Analysis of Vaccination Issues and Sentiments Toward Vaccination

We collected 40,359 posts on childhood vaccination. Vaccination issues were grouped into 17 second-level classes of ontology. Table 1 shows frequencies of each vaccination issue in the posts. *Parents' health belief* appeared in 15,709 out of 29,521 posts (53.21%) and *experience of vaccine administration place* appeared in 14,964 out of 29,521 posts (50.69%). The lowest frequency vaccination issue was *vaccination intention*, which appeared in 198 out of 29,521 posts (0.67%). Sentiments toward vaccination appeared in 27,236 out of 40,359 posts (67.48%). Positive sentiments appeared in 17,454 out of 27,236 posts (64.08%), negative sentiments appeared in 7121 out of 27,236 posts (26.15%), and neutral sentiments appeared in 2661 out of 27,236 posts (9.77%).

XSL•FO
RenderX

Table 1. Frequency of vaccination issues (total posts, N=29,521).

Vaccination issues	n (%)
Parents' health belief	15,709 (53.2)
Experience of vaccine administration place	14,964 (50.7)
Epidemic	13,386 (45.3)
Experience of adverse event	10,125 (34.3)
Medical service use	9002 (30.5)
Child's illness	8512 (28.8)
Type of vaccine	8191 (27.7)
Vaccination availability	8011 (27.1)
Knowledge of vaccination	6774 (22.9)
Experience of vaccination team	5310 (18.0)
Media coverage	5050 (17.1)
Vaccine dose number	5037 (17.1)
Vaccine- and vaccination-related organizations	4886 (16.6)
Vaccination policy	4235 (14.3)
Vaccine administration site	2077 (7.0)
Vaccine administration time	753 (2.6)
Vaccination intention	198 (0.7)

Trend Analysis of Sentiments Toward Vaccination

Monthly sentiment trends with vaccination-related news are shown in Figure 2. Positive sentiments increased in April and September, and negative sentiments increased in February and June. Neutral sentiments showed little variation. In February, when negative sentiments increased, the news about the spread of measles in the United States had been reported. In April, when positive sentiments increased, the news about vaccination weekly events had been reported. In June, when negative sentiments increased, the news about the spread of Hong Kong flu and associated deaths had been reported. In September, when positive sentiments increased, the news about the KCDC recommendations on seasonal influenza vaccinations had been reported.

Sentiment Analysis

The results of multivariate analysis of vaccination issues on sentiments over time are shown in Table 2. Before vaccination, *vaccination policy*, *parents' health belief*, and *vaccination availability* affected positive sentiments; *child's illness* and *knowledge of vaccination* affected negative sentiments. The time of vaccination, *vaccine administration time*, *vaccine dose number*, and *vaccine administration site*, affected positive sentiments. After vaccination, *experience of vaccine adverse event* affected negative sentiments and *experience of vaccine administration place* affected positive sentiments. The results of a multivariate analysis of vaccination issues on sentiments are shown in Table 3. *Vaccine administration time*, *vaccination policy*, *parents' health belief*, *vaccination availability*, *vaccine dose number*, and *vaccine administration site* affected positive sentiments. *Experience of vaccine adverse event*, *child's illness*, and *knowledge of vaccination* affected negative sentiments.

XSL•FO
RenderX

Figure 2. Monthly sentiment trends and vaccination related news.

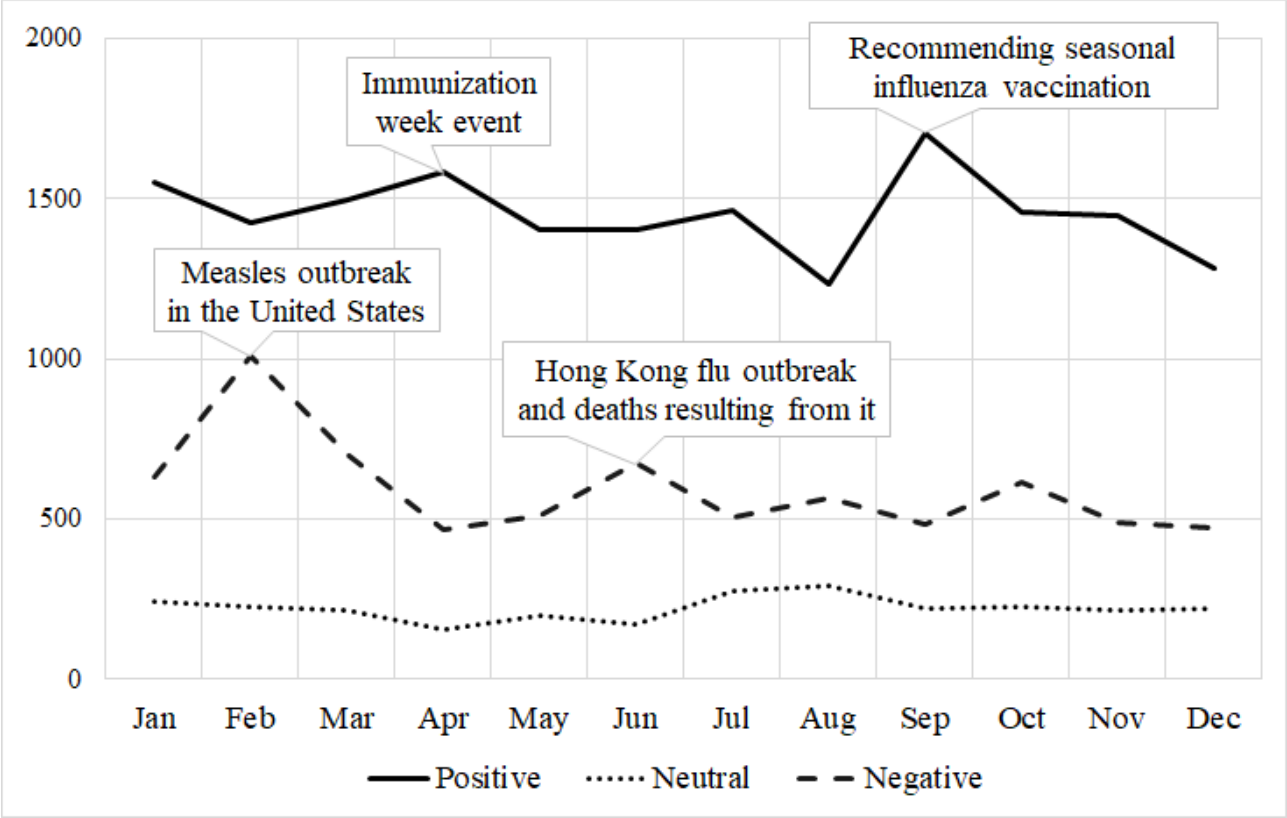


Table 2. Multivariate analyses of vaccination issues on sentiment over time.

Time and vaccination issues	Estimate	SE	z value	*P* value	Odds ratio
Before vaccination					
Vaccination policy	0.80	0.05	15.22	<.001	2.23
Parents' health belief	0.65	0.03	19.63	<.001	1.91
Child's illness	-0.37	0.03	-10.74	<.001	0.69
Vaccination availability	0.29	0.04	8.02	<.001	1.34
Knowledge of vaccination	-0.25	0.04	-6.89	<.001	0.78
Vaccination intention	0.19	0.21	0.92	.36	1.21
Vaccine- and vaccination-related organizations	-0.03	0.04	-0.61	.54	0.98
Vaccination					
Vaccine administration time	0.87	0.13	6.88	<.001	2.38
Vaccine dose number	0.38	0.04	9.23	<.001	1.46
Vaccine administration site	0.17	0.06	2.98	<.001	1.19
After vaccination					
Experience of vaccine adverse event	-0.16	0.03	-5.21	<.001	0.85
Experience of vaccine administration place	0.14	0.03	4.65	<.001	1.15

XSL•FO
RenderX

Table 3. Multivariate analysis of vaccination issues on sentiment.

Vaccination issues	Estimate	SE	z value	*P* value	Odds ratio
Vaccinated administration time	0.72	0.13	5.62	<.001	2.06
Vaccination policy	0.72	0.05	13.41	<.001	2.04
Parents' health belief	0.70	0.03	20.31	<.001	2.01
Experience of vaccine adverse event	-0.33	0.04	-8.53	<.001	0.72
Vaccination availability	0.30	0.04	7.89	<.001	1.35
Vaccine dose number	0.29	0.04	6.84	<.001	1.34
Child's illness	-0.27	0.04	-6.81	<.001	0.76
Vaccine administration site	0.24	0.06	3.93	<.001	1.27
Knowledge of vaccination	-0.23	0.04	-6.25	<.001	0.79
Experience of vaccine administration place	-0.04	0.03	-1.33	.18	0.96
Vaccination intention	0.03	0.21	0.16	.88	1.03
Vaccine- and vaccination-related organizations	0.00	0.04	-0.02	.99	1.00

Table 4. Association rules with top 5 lift and bottom 5 lift.

Rules	Support	Confidence	Lift
Health belief[a], availability[b], policy[c] ⇒ Positive sentiment	0.06	0.92	1.29
Availability, policy, place[d] ⇒ Positive sentiment	0.05	0.90	1.27
Availability, policy ⇒ Positive sentiment	0.07	0.89	1.25
Availability, place, organizations[e] ⇒ Positive sentiment	0.06	0.89	1.25
Health belief, policy ⇒ Positive sentiment	0.09	0.89	1.25
Health belief, knowledge[f], place, AE[g] ⇒ Positive sentiment	0.05	0.57	0.80
Knowledge, place, AE ⇒ Positive sentiment	0.06	0.57	0.81
Health belief, illness[h], AE ⇒ Positive sentiment	0.06	0.59	0.83
Health belief, knowledge, AE ⇒ Positive sentiment	0.07	0.59	0.83
Knowledge, AE ⇒ Positive sentiment	0.07	0.60	0.84

[a]Parents' health belief.

[b]Vaccination availability.

[c]Vaccination policy.

[d]Experience of vaccine administration place.

[e]Vaccine- and vaccination-related organizations.

[f]Knowledge of vaccination.

[g]AE: experience of vaccine adverse event.

[h]Child's illness.

The results of analysis of sets of vaccination issues associated with sentiments are shown in Table 4.

Of the 80 rules, Table 4 shows the 5 rules with the highest lift and the 5 rules with the lowest lift. For example, the sets of *parents' health belief*, *vaccination availability*, and *vaccination policy* were associated with positive sentiments with support of 0.06, confidence of 0.92, and lift of 1.29. A support level of 0.06 indicates that 6% of posts have *parents' health belief*, *vaccination availability*, and *vaccination policy* along with positive sentiments. A confidence level of 0.92 indicates that 92% of posts with *parents' health belief*, *vaccination availability*, and *vaccination policy* have positive sentiments. A lift level of 1.29 indicates that the ratio of appearance of positive sentiments in posts with *parents' health belief*, *vaccination availability*, and *vaccination policy* to the appearance of positive sentiments in total posts is 1.29. The set of *parents' health belief*, *knowledge of vaccination*, *experience of vaccine administration place*, and *experience of vaccine adverse event* was associated with positive sentiments with a lift of 0.80. This indicates that the ratio of appearance of positive sentiments in posts with *parents' health belief*, *knowledge of vaccination*, *experience of vaccine administration place*, and *experience of vaccine adverse event* to the appearance of positive sentiments in total posts is 0.80. In other words, the ratio of

XSL•FO
RenderX

appearance of negative sentiments in posts with these vaccination issues to the appearance of negative sentiments in total posts is 1.25.

Discussion

Principal Findings

We developed a childhood vaccination ontology as a framework for systematically collecting and analyzing social data on childhood vaccination and used this ontology to identify public concerns about and sentiments toward childhood vaccination from social data.

The childhood vaccination ontology developed in this study had the following characteristics. First, this ontology was the first ontology describing childhood vaccination, including factors affecting vaccination, as well as vaccination intention, behavior, and experience. Although there was a preexisting Vaccine Ontology [29] available, it was not suitable for identifying public concerns about and sentiments toward childhood vaccination from social data. Second, this ontology included various factors affecting vaccination, such as *child*, *parent*, *family and friend*, *organization and institution*, *society and community*, and *policy*, by modifying the Social Ecological Framework [32]. This ontology was not limited to factors affecting vaccination but also included intention before vaccination, vaccination behavior, and experience after vaccination. Third, this ontology included Is-A and attribute relationships among concepts. Classes were modeled as EAVs and thus included attributes of each class and values of the attributes. Finally, this ontology included terminology with synonyms for classes, attributes, and values of attributes so that we can collect and analyze the social media posts on childhood vaccination.

According to frequency analysis, the most common vaccination issue was *parent's health belief* (53.2%). This was also one of the important topics in other surveys on vaccination decision making and hesitancy [5,7,8,12,30,31,39]. Positive sentiments appeared 2.5 times more than negative sentiments. Positive sentiments toward vaccination were also identified more often than negative sentiments in other studies based on social data [14,18,40]. Members of the public who consider vaccination to be a useful measure to prevent infectious diseases might post narratives with positive sentiments.

According to a trend analysis of sentiments, public sentiments toward vaccination fluctuated with news about vaccination. Positive sentiments increased when news encouraging vaccination, such as vaccination campaigns, was announced, whereas negative sentiments increased when news about epidemic outbreaks, such as the measles outbreak in the United States, was announced. Other studies have also identified increases in vaccine-related posts, including positive or negative sentiments, to news stories about vaccination [41,42].

According to logistic regression analysis, the public who are aware of aspects of vaccination policy, such as promotion of free vaccinations and the increasing number of medical institutions offering free vaccinations, and public with health belief that vaccination is preventing infectious disease and such diseases are serious if not prevented, posted positive sentiments whereas the public who experienced unwanted adverse events after vaccination posted negative sentiments. Therefore, it is important to inform the public of vaccination policy, the benefits of vaccination, and how to deal with adverse events to lower negative sentiments, which can affect vaccination intention or behavior.

According to association analysis of vaccination issues with sentiment, the public who posted vaccination policy and vaccination availability in terms of cost and distance to appropriate medical institutions posted more positive sentiments whereas the public who posted adverse event experienced after vaccination and knowledge of vaccination posted more negative sentiments. This was also found in logistic regression analysis. *Parents' health belief* and *experience of vaccine administration place* were associated with both positive and negative sentiments. These were posted in more than half of analyzed posts in which both positive and negative sentiments appeared.

Limitations

We classified the sentiments of the posts by comparing the numbers of words expressing positive and negative emotions. We did not reflect weight and degree of emotion, such as severe and mild, and did not distinguish double negatives or tense of words when classifying sentiments. There are other studies used machine learning algorithms [17,42] to classify the emotions in social data or a sentiment score [19] that expresses sentiments as a numerical value. We suggest developing and using new emotion classification algorithm reflecting weight and degree of emotion or sentiment score as further research.

We were not able to identify yearly trends in sentiments and vaccination issues because of the short duration of data collection period. We suggest studying yearly sentiments trends using data collected from more than 1 year in future research.

We were not able to study vaccination intention or behavior because of a lack of posts discussing these topics. Vaccination intention and behavior can be studied by combining various data sources, such as survey data and existing vaccination statistics from the immunization registry in the future.

We used only those vaccination issues classified in terms of the second-level class concepts of the ontology developed in this study. We suggest applying lower-level class concepts and their relationships, for example, different types of vaccination and types of adverse events, in future research.

Use of social data for research is justified because of public accessibility of social media data. However, there exist ethical concerns around privacy and the protection of sensitive information, because it involves collecting data involving human subjects [43]. The public might stop posting their concerns or opinions on vaccination-related issues on social media if they realize that their posts are being analyzed by data scientists. To solve these issues, we anonymized social data by removing identifying information and did not use any social media quotes that might identify a social media user.

XSL•FO
RenderX

Conclusions

In this study, we developed a childhood vaccination ontology comprising 9 superclasses and 124 subclasses with 4 levels of depth and a terminology containing 882 synonyms for class, attribute, and value concepts. We used this ontology as a framework to identify the public concerns about and sentiments toward childhood vaccination from social data. This is the first study to analyze public concerns about and sentiments toward childhood vaccination using social media posts by developing an ontology.

Parent's health belief, *vaccination availability*, and *vaccination policy* were the 3 most significant factors associated with positive sentiment. Health belief may be influenced by antivaccine arguments such as the view that natural immunity is better than vaccine-acquired immunity or baseless rumors claiming that vaccines cause autism. Thus, it is important to monitor antivaccine arguments and rumors posted on the social media that might increase negative sentiment toward vaccination. *Vaccination availability* including cost and the travel distance to vaccine administration place is related to *vaccination policy*, such as the increasing number of free vaccines and number of health care institutions offering free vaccinations. Thus, it is important to publicize policies on free vaccinations to improve positive sentiments toward vaccination. As negative sentiments toward vaccination affect people's intention to vaccinate and thus lead to a reduction in vaccination rates [5-9], it is important to introduce ways to improve positive sentiments toward vaccination.

We expect that practitioners and researchers in the field of childhood vaccination may use this ontology to identify public concerns about and sentiments toward childhood vaccination from social data.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2018R1A2A2A05022021).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Social media channels: a list of social media channels used to collect social data.

[XLSX File (Microsoft Excel File), 13KB - jmir_v21i6e13456_app1.xlsx]

References

1. Disease control center. Chungbuk: KCDC; 2017. Epidemiology and management of vaccine preventable disease (5th ed) URL: http://www.cdc.go.kr/CDC/together/CdcKrTogether0302.jsp?menuIds=HOME001-MNU1154-MNU0725-MNU0088&fid=10713&q_type=&q_value=&cid=138079 [accessed 2018-03-11] [WebCite Cache ID 71ydiQtRo]
2. KCDC. Disease control center. Internet website of vaccination assistance URL: https://nip.cdc.go.kr/irgd/index.html [accessed 2018-03-11] [WebCite Cache ID 71yg1Ggfc]
3. KCDC. Public Data Portal. Chungbuk: KCDC; 2017. National Childhood Vaccination Coverage Among Children Aged 1-3 years in Korea URL: https://www.data.go.kr/dataset/3075043/fileData.do [accessed 2018-03-11] [WebCite Cache ID 75SShEdeQ]
4. Dubé E, Vivion M, MacDonald NE. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. Expert Rev Vaccines 2015 Jan;14(1):99-117. [doi: 10.1586/14760584.2015.964212] [Medline: 25373435]
5. Kim SK. RISS Mobile. Seoul, Korea: Chung-Ang University; 2007. Configuration of a vaccination-behavior model for mothers with infant URL: http://www.riss.kr/search/detail/DetailView.do?p_mat_type=be54d9b8bc7cdb09&control_no=0170e0599bd183d9ffe0bdc3ef48d419 [accessed 2018-03-11] [WebCite Cache ID 71yk095OD]
6. Cha HG, Ryoo EN, Park SH. Factors affecting the intention of vaccination in parents with noncompliance of vaccination. Korean J Health Educ Promot 2012;29(1):89-96 [FREE Full text]
7. Kim KS, Kim BJ, Kim KH, Kwon SH, Kim SK. Factors associated with intention for additional vaccination in mothers with children aged 4~6 Years. J Korean Acad Child Health Nurs 2007;13(4):478-485 [FREE Full text]
8. Yim ES, Lee KJ, Cheon EY, Lim MR. Factors influencing children's immunization. J Korean Acad Community Health Nurs 2006;17(2):283-294 [FREE Full text]
9. Gust DA, Darling N, Kennedy A, Schwartz B. Parents with doubts about vaccines: which vaccines and reasons why. Pediatrics 2008 Oct;122(4):718-725 [FREE Full text] [doi: 10.1542/peds.2007-0538] [Medline: 18829793]
10. Kennedy A, Lavail K, Nowak G, Basket M, Landry S. Confidence about vaccines in the United States: understanding parents' perceptions. Health Aff (Millwood) 2011 Jun;30(6):1151-1159 [FREE Full text] [doi: 10.1377/hlthaff.2011.0396] [Medline: 21653969]

XSL•FO
RenderX

11. Rhim JW, Kim CH, Lee WB, Kang JH. A survey of parental knowledge of vaccination. Korean J Pediatr 2006;49(3):251-257 [FREE Full text]
12. Ko JY, Choi YS, Hong JI. The trend in the public's awareness and policy satisfaction on immunization in Korean. Public Health Rep 2015;8(28):638-647 [FREE Full text]
13. Song TM, Ryu S. Big data analysis framework for healthcare and social sectors in Korea. Healthc Inform Res 2015 Jan;21(1):3-9 [FREE Full text] [doi: 10.4258/hir.2015.21.1.3] [Medline: 25705552]
14. Kang GJ, Ewing-Nelson SR, Mackey L, Schlitt JT, Marathe A, Abbas KM, et al. Semantic network analysis of vaccine sentiment in online social media. Vaccine 2017 Dec 22;35(29):3621-3638 [FREE Full text] [doi: 10.1016/j.vaccine.2017.05.052] [Medline: 28554500]
15. UNICEF. New York; 2013. Tracking anti-vaccination sentiment in Eastern European social media networks URL: https://www.unicef.org/eca/reports/tracking-anti-vaccination-sentiment-eastern-european-social-media-networks [accessed 2019-01-22] [WebCite Cache ID 75c3nJjqW]
16. UN Global Pulse. Understanding public perceptions of immunisation using social media. Glob Pulse Proj Ser 2014;9:1-2 [FREE Full text]
17. Deiner MS, Fathy C, Kim J, Niemeyer K, Ramirez D, Ackley SF, et al. Facebook and Twitter vaccine sentiment in response to measles outbreaks. Health Informatics J 2017 Nov 1:1460458217740723. [doi: 10.1177/1460458217740723] [Medline: 29148313]
18. Salathé M, Khandelwal S. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. PLoS Comput Biol 2011 Oct;7(10):e1002199 [FREE Full text] [doi: 10.1371/journal.pcbi.1002199] [Medline: 22022249]
19. Keim-Malpass J, Mitchell EM, Sun E, Kennedy C. Using Twitter to understand public perceptions regarding the# HPV vaccine: opportunities for public health nurses to engage in social marketing. Public Health Nurs 2017 Dec;34(4):316-323. [doi: 10.1111/phn.12318] [Medline: 28261846]
20. Ministry of Science and ICT, Korea Internet & security agency. Korea Internet & Security Agency. Jeollanam-do, Korea; 2018. 2017 Survey on the internet usage URL: http://www.kisa.or.kr/eng/usefulreport/surveyReport_View.jsp?mode=view&p_No=4&b_No=262&d_No=81 [accessed 2018-12-13] [WebCite Cache ID 75c3vOJhU]
21. Kim SH. dCollection Digital Academic Information Distribution System. Seoul, Korea: Graduate School of Sungkyunkwan University; 2015. Developing an ontology and its usability to social data analytics URL: http://www.dcollection.net/handler/skku/000000080135 [accessed 2018-03-11] [WebCite Cache ID 75Yohu4ew]
22. Yu E, Kim J, Lee C, Kim N. Using ontologies for semantic text mining. J Inf Syst 2012 Sep 30;21(3):137-161. [doi: 10.5859/KAIS.2012.21.3.137]
23. Kim AR, Park H, Song T. Development and evaluation of an obesity ontology for social big data analysis. Healthc Inform Res 2017 Jul;23(3):159-168 [FREE Full text] [doi: 10.4258/hir.2017.23.3.159] [Medline: 28875050]
24. Jung H, Park H, Song T. Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals. J Med Internet Res 2017 Jul 24;19(7):e259 [FREE Full text] [doi: 10.2196/jmir.7452] [Medline: 28739560]
25. Bae JY. Ewha Womans University. Seoul, Korea: Ewha Womans University; 2016. Prediction on Food Safety Risk Attached to Social Big Data Analysis Based on Ontology URL: http://dspace.ewha.ac.kr/handle/2015.oak/214711 [accessed 2018-03-11] [WebCite Cache ID 75asfUjJ9]
26. Noy NF, McGuinness DL. Stanford University. 2001. A Ontology development 101: A guide to creating your first ontology URL: https://protege.stanford.edu/publications/ontology_development/ontology101.pdf [accessed 2018-03-11] [WebCite Cache ID 75Z9GmwNs]
27. Grüninger M, Fox MS. Methodology for the design and evaluation of ontologies. In: Workshop Notes of IJCAI-95. Montreal, Canada; 1995 Presented at: Workshop on Basic Ontological Issues in Knowledge Sharing; April 7, 1995; Montreal, Canada p. 1-10 URL: http://stl.mie.utoronto.ca/publications/gruninger-ijcai95.pdf
28. National Center for Biomedical Ontology. URL: https://bioportal.bioontology.org/ [accessed 2019-01-03] [WebCite Cache ID 77lyt9dzG]
29. He Y, Chen F, Scheuermann R, Todd T, Brinkman R, Colby L, et al. VO: Vaccine Ontology. 2009 Aug 5 Presented at: The 1st International Conference on Biomedical Ontology (ICBO 2009); July 24-26, 2009; Buffalo, NY, USA. [doi: 10.1038/npre.2009.3552.1]
30. Sturm LA, Mays RM, Zimet GD. Parental beliefs and decision making about child and adolescent immunization: from polio to sexually transmitted infections. J Dev Behav Pediatr 2005 Dec;26(6):441-452. [Medline: 16344662]
31. Katz IT, Ware NC, Gray G, Haberer JE, Mellins CA, Bangsberg DR. Scaling up human papillomavirus vaccination: a conceptual framework of vaccine adherence. Sex Health 2010 Sep;7(3):279-286 [FREE Full text] [doi: 10.1071/SH09130] [Medline: 20719215]
32. Kumar S, Quinn SC, Kim KH, Musa D, Hilyard KM, Freimuth VS. The social ecological model as a framework for determinants of 2009 H1N1 influenza vaccine uptake in the United States. Health Educ Behav 2012 Apr;39(2):229-243 [FREE Full text] [doi: 10.1177/1090198111415105] [Medline: 21984692]

XSL•FO
RenderX

33. Larson HJ, Jarrett C, Eckersberger E, Smith DM, Paterson P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012. Vaccine 2014 Apr 17;32(19):2150-2159. [doi: 10.1016/j.vaccine.2014.01.081] [Medline: 24598724]
34. Dubé E, Laberge C, Guay M, Bramadat P, Roy R, Bettinger J. Vaccine hesitancy: an overview. Hum Vaccin Immunother 2013 Aug;9(8):1763-1773 [FREE Full text] [doi: 10.4161/hv.24657] [Medline: 23584253]
35. KCDC. Disease control center. Chungbuk: KCDC; 2013. Epidemiology and management of vaccine preventable disease. 4th ed URL: http://www.cdc.go.kr/CDC/together/CdcKrTogether0302.jsp?menuIds=HOME006-MNU2804-MNU3027-MNU2979&cid=138282 [accessed 2016-03-25] [WebCite Cache ID 75c4n3INA]
36. Centers for Disease Control and Prevention. Washington D.C: Public Health Foundation; 2015. Epidemiology and Prevention of Vaccine-Preventable Diseases. 13th ed URL: https://www.cdc.gov/vaccines/pubs/pinkbook/index.html [accessed 2016-03-25] [WebCite Cache ID 75c4AiB6j]
37. Public Health England. GOV.uk. London: Public Health England; 2013. Immunisation against infectious disease URL: https://www.gov.uk/government/collections/immunisation-against-infectious-disease-the-green-book#the-green-book [accessed 2016-03-25] [WebCite Cache ID 75c4EFXOe]
38. SK telecom. Smart Insight by SK telecom. URL: http://www.smartinsight.co.kr/ [accessed 2019-05-28] [WebCite Cache ID 78hTLBjNe]
39. Oh HK. Seoul National University. Seoul, Korea: Seoul National University; 2015. Factors associated with vaccination of the 4th dose of DTaP in children aged 3 years URL: http://s-space.snu.ac.kr/handle/10371/124846 [accessed 2018-03-11] [WebCite Cache ID 75azDvK4g]
40. Keelan J, Pavri V, Balakrishnan R, Wilson K. An analysis of the Human Papilloma Virus vaccine debate on MySpace blogs. Vaccine 2010 Feb 10;28(6):1535-1540. [doi: 10.1016/j.vaccine.2009.11.060] [Medline: 20003922]
41. Dredze M, Broniatowski DA, Smith MC, Hilyard KM. Understanding vaccine refusal: why we need social media now. Am J Prev Med 2016 Apr;50(4):550-552. [doi: 10.1016/j.amepre.2015.10.002] [Medline: 26655067]
42. Tomeny TS, Vargo CJ, El-Toukhy S. Geographic and demographic correlates of autism-related anti-vaccine beliefs on Twitter, 2009-15. Soc Sci Med 2017 Dec;191:168-175 [FREE Full text] [doi: 10.1016/j.socscimed.2017.08.041] [Medline: 28926775]
43. Townsend L, Wallace C. The ethics of using social media data in research: a new framework. In: Woodfield K, editor. The Ethics of Online Research (Advances in Research Ethics and Integrity, Volume 2). Bingley: Emerald Publishing Limited; 2017:189-207.

Abbreviations

CVI: content validity index
DL: description logic
EAV: entity-attribute-value
HPV: Human Papilloma Virus
KCDC: Korean Centers for Disease Control and Prevention
MMR: Measles, Mumps, and Rubella
NRF: National Research Foundation of Korea
SNS: Social Networking Service

Edited by G Eysenbach; submitted 22.01.19; peer-reviewed by W Erdley, J Brixey, R Bartlett; comments to author 14.03.19; revised version received 09.05.19; accepted 10.05.19; published 07.06.19.

Please cite as:
On J, Park HA, Song TM
Sentiment Analysis of Social Media on Childhood Vaccination: Development of an Ontology
J Med Internet Res 2019;21(6):e13456
URL: http://www.jmir.org/2019/6/e13456/
doi:10.2196/13456
PMID:31199290

©Jeongah On, Hyeoun-Ae Park, Tae-Min Song. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete

XSL•FO
RenderX

bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Detecting Signs of Depression in Tweets in Spanish: Behavioral and Linguistic Analysis

Angela Leis[1], PsyM; Francesco Ronzano[1], PhD; Miguel A Mayer[1], MD, PhD; Laura I Furlong[1], PhD; Ferran Sanz[1], Prof Dr

Research Programme on Biomedical Informatics, Hospital del Mar Medical Research Institute, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, Barcelona, Spain

Corresponding Author:
Ferran Sanz, Prof Dr
Research Programme on Biomedical Informatics
Hospital del Mar Medical Research Institute
Department of Experimental and Health Sciences, Universitat Pompeu Fabra
Carrer Dr Aiguader 88
Barcelona, 08003
Spain
Phone: 34 933 160 540
Fax: 34 933 160 550
Email: ferran.sanz@upf.edu

Abstract

Background: Mental disorders have become a major concern in public health, and they are one of the main causes of the overall disease burden worldwide. Social media platforms allow us to observe the activities, thoughts, and feelings of people's daily lives, including those of patients suffering from mental disorders. There are studies that have analyzed the influence of mental disorders, including depression, in the behavior of social media users, but they have been usually focused on messages written in English.

Objective: The study aimed to identify the linguistic features of tweets in Spanish and the behavioral patterns of Twitter users who generate them, which could suggest signs of depression.

Methods: This study was developed in 2 steps. In the first step, the selection of users and the compilation of tweets were performed. A total of 3 datasets of tweets were created, a depressive users dataset (made up of the timeline of 90 users who explicitly mentioned that they suffer from depression), a depressive tweets dataset (a manual selection of tweets from the previous users, which included expressions indicative of depression), and a control dataset (made up of the timeline of 450 randomly selected users). In the second step, the comparison and analysis of the 3 datasets of tweets were carried out.

Results: In comparison with the control dataset, the depressive users are less active in posting tweets, doing it more frequently between 23:00 and 6:00 (P<.001). The percentage of nouns used by the control dataset almost doubles that of the depressive users (P<.001). By contrast, the use of verbs is more common in the depressive users dataset (P<.001). The first-person singular pronoun was by far the most used in the depressive users dataset (80%), and the first- and the second-person plural pronouns were the least frequent (0.4% in both cases), this distribution being different from that of the control dataset (P<.001). Emotions related to sadness, anger, and disgust were more common in the depressive users and depressive tweets datasets, with significant differences when comparing these datasets with the control dataset (P<.001). As for negation words, they were detected in 34% and 46% of tweets in among depressive users and in depressive tweets, respectively, which are significantly different from the control dataset (P<.001). Negative polarity was more frequent in the depressive users (54%) and depressive tweets (65%) datasets than in the control dataset (43.5%; P<.001).

Conclusions: Twitter users who are potentially suffering from depression modify the general characteristics of their language and the way they interact on social media. On the basis of these changes, these users can be monitored and supported, thus introducing new opportunities for studying depression and providing additional health care services to people with this disorder.

(J Med Internet Res 2019;21(6):e14199) doi:10.2196/14199

KEYWORDS

depression; social media; mental health; text mining

XSL•FO
RenderX

Introduction

Background

Mental health is an essential component of our health. The World Health Organization (WHO) defines mental health as a "state of well-being in which people realize their potential, cope with the normal stresses of life, work productively, and contribute to their communities" [1]. Good mental health is about being cognitive, emotionally and socially healthy and it helps to determine the way we think and feel, in relation with others and how we make choices. Several factors, such as genetic, sociocultural, economic, political and environmental aspects, shape and influence our mental health. In the last few years, mental disorders have become a major concern in public health, and they are one of the main causes of the overall disease burden worldwide. They have devastating consequences for both patients and their families [2-7]. According to the WHO, depressive disorders are the most common among the mental illnesses [8]. Such disorders conditions are characterized by sadness, loss of interest and pleasure, feelings of guilt or low self-worth, disturbed sleep or appetite, feelings of tiredness, and poor concentration [8]. In 2018, at the global level, more than 300 million people were suffering from depression, and it is the main contributor to global disability. Depression has several consequences, both personal and social costs [9,10]. In some cases, depression can lead to suicide ideation and attempts [2,11]. The prevalence of this disorder changes depending on age, but it affects the whole population, from children and adolescents to elderly people. From 2005 to 2015, the number of people with depression increased by around 18% [12]. In this context, social media platforms allow to observe the activities, thoughts, and feelings of people's daily lives and thereby investigate their emotional well-being. This domain has become a new growing area of interest in public health and health care research [13-16]. People with depression often use social media to talk about their illness and treatment, share information and experiences, seek social support and advice, reduce social isolation, and manage their mental illness [15-21]. In addition, access to mobile devices facilitates the use of social media platforms, such as Twitter and Facebook, at any time and at any place. Social media, such as Twitter, is by nature social, and we can consequently find social patterns in Twitter feeds, thereby revealing key aspects of mental and affective disorders [22]. Social media has become an important source of health-related information, which allows us to detect and predict affective disorders and which can be used as an additional tool for mental health monitoring and infoveillance [23-26]. Furthermore, the application of different methodologies based on natural language processing and machine learning technologies has proved to be effective in supporting and automating the identification of early signs of mental illness by analyzing the content shared on the Web by individuals [13-15,27]. This human interaction with social media contributes to build the so-called digital phenotype, reshaping disease expression in terms of the lived experience of individuals and detecting early manifestations of several conditions [28]. Twitter is an internet microblogging social media service that allows users to post short messages about facts, feelings and opinions, and, as shown in previous studies, users' health conditions [15]. Twitter is one of the most important social media platforms in terms of number of users, with more than 330 million active users worldwide [29]. Since November 2017, the maximum number of characters of a tweet has been increased from 140 to 280. By analyzing huge amounts of text, researchers can link everyday language use with social behavior and personality [30,31]. Language, as a means of communication, constitutes an essential element for providing valuable insights about people's interests, feelings and concerns [32]. For this reason, the analysis of the messages posted on social media platforms may provide information about many personality traits, lifestyles, and psychological disorders [13,33,34]. The potential anonymity of social media encourages its users to be more willing to report health information, such as details of their mental disorders and treatments received. In addition, it is seen as a way to communicate and receive support from others with similar experiences, avoiding the isolation and fighting the social stigma of these conditions [12,15,17,19,32,35]. Nevertheless, users suffering from depression may also feel uncomfortable socializing and consuming information on social media platforms [36]. Several features of the messages, such as number and frequency of tweets, distribution throughout the day or during the night hours, and their seasonal character, can be used for the detection and monitoring of mental disorders, such as depression [20]. This knowledge can help health care professionals and health institutions and services in the decision-making processes to ensure better management of patients suffering from depression.

Objectives

There are many studies that have used data mining and machine learning techniques on social media platforms to automatically identify people with mental health problems, such as depression, posttraumatic stress disorder, schizophrenia, or eating disorders, usually focusing the studies on messages written in English [20,37-39]. As far as we know, on social media, there are no studies about mental disorders that analyze messages written in Spanish. Taking into account that Spanish speaking countries, such as Spain and Mexico, are among the 10 most active Twitter users in the world, with more than 6 million and 7 million users, respectively [40], we focused our research on the expression of depression in Spanish language tweets. The aim of this study was to identify the linguistic features of tweets written in Spanish and the behavioral patterns of the corresponding Twitter users that could suggest signs of depression.

Methods

Study Steps

This study was designed and developed in 2 steps, with the aim of analyzing the linguistic patterns and behavioral features of Twitter users suffering from depression in comparison with the general population of Twitter users. The study was focused on tweets written in Spanish. In the first step, the selection of users and the compilation of tweets were performed. Given the design and purpose of the study, we decided to use the Twitter Application Programming Interface (API) [41]. Using this API, 3 datasets of tweets were created:

XSL•FO
RenderX

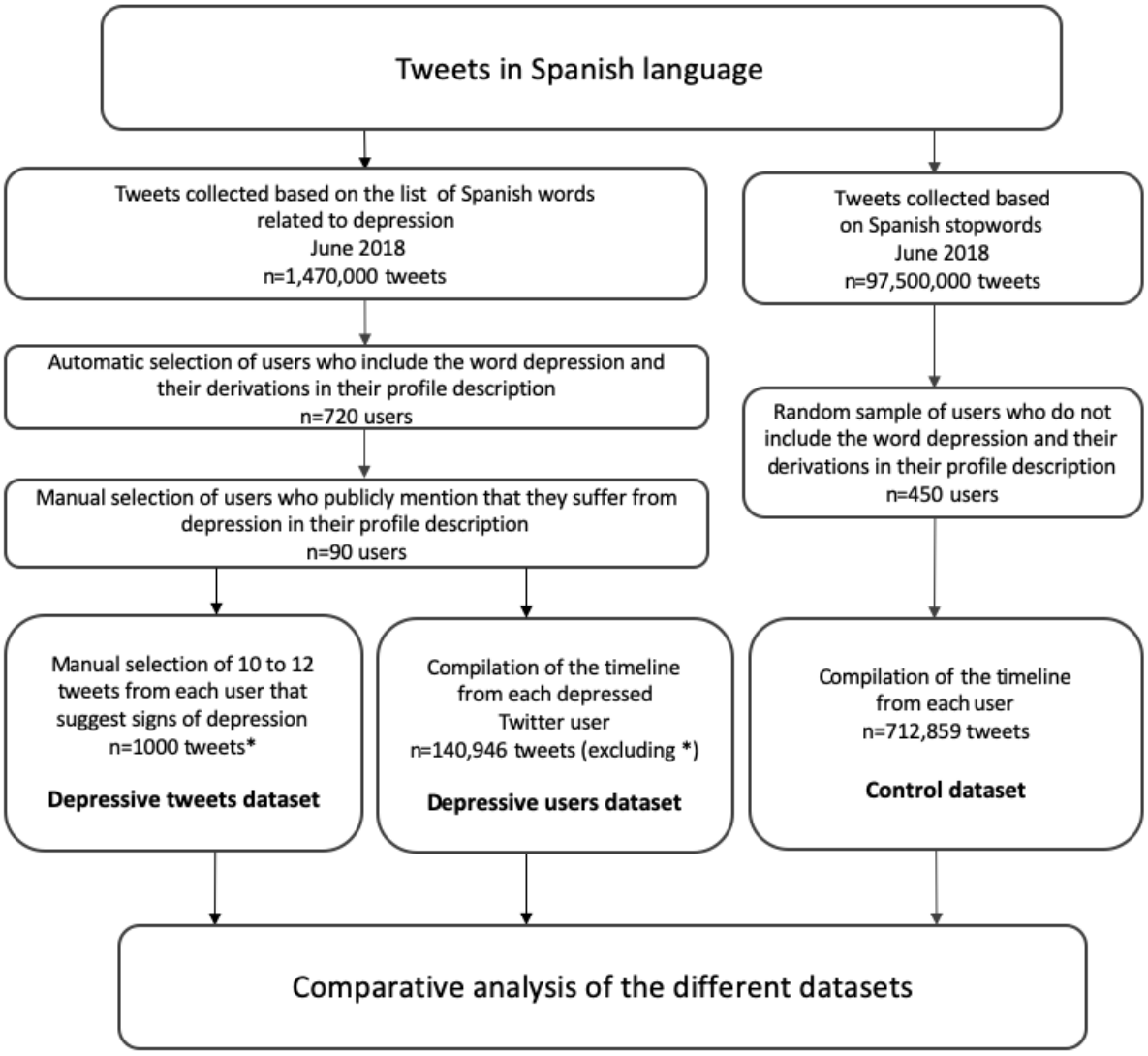
1. The *depressive users* dataset was made up of the timeline of 90 users who publicly mentioned on their Twitter profile that they suffer from depression.
2. The *control* dataset was made up of the timeline of 450 randomly selected Twitter users.
3. The *depressive tweets* dataset was constituted by a manual selection of tweets from the depressive users dataset, which specifically included expressions indicative of depression.

In the second step, comparison and analysis of the 3 datasets of tweets (control, depressive users, and depressive tweets datasets) were carried out to spot their distinguishing features. In the rest of this section, we will describe the methodology in detail. The flow diagram of the study is depicted in Figure 1.

Data Collection and User Selection

The selection of the tweets and their users was based on the filtered real-time streaming support provided by the Twitter API. In the first step, we selected the users who showed potential signs of depression on Twitter on the basis of the 20 most frequent words in Spanish expressed by patients suffering from depression in clinical settings. These words were jointly identified and selected by a psychologist and a family physician with clinical experience and were based on the definition and general features of depression according to the Diagnostic and Statistical Manual of Mental Disorders [42]. The list of words used and their English translations are shown in Textbox 1.

Figure 1. Flow diagram of the study process.



XSL•FO
RenderX

Textbox 1. List of Spanish words related to depression and their English translations.

- agobiado/a (overwhelmed)
- agotado/a (exhausted)
- angustiado/a (distressed)
- ansiedad (anxiety)
- ansioso/a (anxious)
- cansado/a (tired)
- decaído (low)
- depresión (depression)
- depresivo/a (depressed as a condition)
- deprimido/a (depressed as state)
- desanimado/a (discouraged)
- desesperado/a (desperate)
- desmotivado/a (demotivated)
- insomnio (insomnia)
- llorar (cry)
- nervioso (nervous)
- preocupado/a (worried)
- solo/a (lonely)
- triste (sad)
- vacío/a (empty)

During June 2018, 1,470,000 tweets, including 1 or more occurrences of the words listed in Textbox 1, were collected. From this collection of tweets and to select the users who publicly stated in the textual description associated to their profile that they suffered from depression, all the profile descriptions, including 1 or more occurrences of the word "depr" and all the possible derivations related to the word depression in Spanish, such as "depre," "depresión," "depresivo," "depresiva," "deprimido," and "deprimida," were considered. From the 720 users who included 1 or more of these words in their description profile, 90 users who stated they suffered from depression or were receiving treatment for depression were selected for the analysis. This selection was performed by a psychologist, verifying that the statements were related to real expressions of depression, excluding quotes, jokes, or fake ones. For each of these depressed Twitter users, we collected all the most recent tweets from their timeline, up to a maximum of about 3200 tweets. Thus, a total of 189,669 tweets were collected, a figure that was reduced to 140,946 after discarding the retweets. These 140,946 tweets constituted the *depressive users dataset*. Examples of sentences appearing in the user profiles that were used for selecting the depressive users are:

- "Paciente psiquiátrico con depresión crónica" (*Psychiatric patient with chronic depression;* example of a profile sentence that indicates depression).
- "Colecciono errores traducidos a tweets depresivos y a uno que otro impulso de amor" (*I gather errors translated into depressing tweets and into one or another love impulse*; example of a profile sentence that does not indicate depression).

Once the users with profile sentences indicating depression had been retrieved, their Twitter timelines were collected. Only those users having in their timeline at least 10 tweets that suggested signs of depression were retained for further analyses. For each user, the selection of these tweets was performed by manually inspecting the tweets of the user's complete timeline in reverse temporal order, starting from the most recent one to the oldest tweet of the timeline retrieved by means of the Twitter API . Finally, a total number of 1000 tweets issued by the 90 depressive users, suggesting signs of depression, were detected and used for the analysis. This set of tweets provided us with the *depressive tweets dataset*, which was used to analyze linguistic features of tweets showing signs of depression. It has to be mentioned that these 1000 tweets were not to be included in the depressive users dataset (see Figure 1). At the same time, more than 97,500,000 tweets were also collected in June 2018: such tweets were gathered by listening to the public Twitter stream during this time span by only considering tweets with Spanish textual contents (as detected by Twitter language identification support).

Given that Twitter requires more restrictive filters than just the language of the tweets, we used a list of the most frequently used Spanish words (stopwords) to retrieve all tweets that included 1 or more of these words. The vast majority of Spanish tweets should match this criterion. A sample of 450 users who did not mention in their profile the word depression and its

XSL•FO
RenderX

derivations were selected randomly from the 97,500,000 tweets. The complete timelines of these users were compiled (1,141,021 tweets), which were reduced to 712,589 once retweets were removed. These 712,589 tweets constituted the *control dataset*. To identify the language of a tweet, we relied on the language automatically identified by Twitter for each tweet, selecting tweets in Spanish. It has to be noted that these data can contain some tweets from unidentified depressive users.

Data Analysis

A comparison of the 3 datasets was performed to determine the existence of differential linguistic and behavioral features. The different features that were analyzed are shown in Table 1.

The textual content of each tweet was analyzed by means of the following sequence of steps:

- Tokenization performed by means of a custom Twitter tokenizer included in the Natural Language Toolkit [43].
- Part-of-Speech (POS) tagging performed by means of the Freeling Natural Language Processing tool in order to analyse the usage patterns of grammatical categories (eg, adjectives, nouns, or pronouns) in the text of tweets [44].
- Identification of negations performed by relying on a custom list of Spanish negation expressions, such as *nada* (nothing), *nadie* (nobody), *no* (no), *nunca* (never), and alike.
- Identification of occurrences of positive and negative words inside the text of each tweet by means of 2 Spanish polarity lexicons: the Spanish Sentiment Lexicon and the Spanish SentiCon Lexicon [45,46]. We exploited 2 lexicons to consider and compare 2 approaches of modeling polarity in Spanish texts, thus reducing any language modeling bias that the use of a single language resource could introduce.
- Identification of words and expressions associated to the basic emotions [47] by using the Spanish Emotion Lexicon [48]. Such emotions are *alegría* (happiness), *enojo* (anger), *miedo* (fear), *repulsión* (disgust), *sorpresa* (surprise), and *tristeza* (sadness).

All the tools and aforementioned resources are publicly available. The statistical analyses were carried out with the R version 3.4.3 (R Development Core Team) and SPSS Statistics version 23.0 (IBM), applying the relevant test for each type of comparison to be carried out.

Table 1. Characteristics of the tweets analyzed.

Feature	Analyses performed
Distributions over time	• Tweets throughout the day (per hour) • Tweets throughout the week
Part-of-Speech	• Number of words by grammatical categories (part-of-speech tags) • Number of personal pronouns
Counts	• Number of characters • 200 most frequent words (word cloud) • Number of hashtags, links, mentions, and emojis
Emotion analysis	• Emotion types and frequencies
Negations	• Negation words types and frequencies
Polarity analyses	• Polarity of tweets on the basis of Spanish Sentiment Lexicon and Spanish SentiCon Polarity

Ethical Approval

The protocol used in this study was approved by the Ethics Committee of Parc Salut Mar (approval number 2017/7234/1).

Results

Distribution Over Time

Regarding the distribution of tweets over time, the number of tweets per hour and throughout the week of control and depressive users datasets were compared. The tweet hours were adjusted by the user's time zone. As shown in Figure 2, the depressive users are less active in generating tweets than the control ones, reaching both groups the same activity level between 23:00 and 6:00. The comparison of the temporal distributions of tweets between both datasets was carried out by means of a repeated measures analysis of variance (Greenhouse-Geisser F=6.605; $P<.001$). As shown in Figure 3, the activity throughout the week of the depressive users dataset presented more regular activity than the control dataset, whose users' activity showed a sharp drop during the weekend. The differences between these datasets were statistically significant (Greenhouse-Geisser F=4.153; $P=.008$).

XSL•FO
RenderX

Figure 2. Number of tweets and retweets per hour of the control and depressive users datasets (mean±standard error of mean). SEM: standard error of mean.

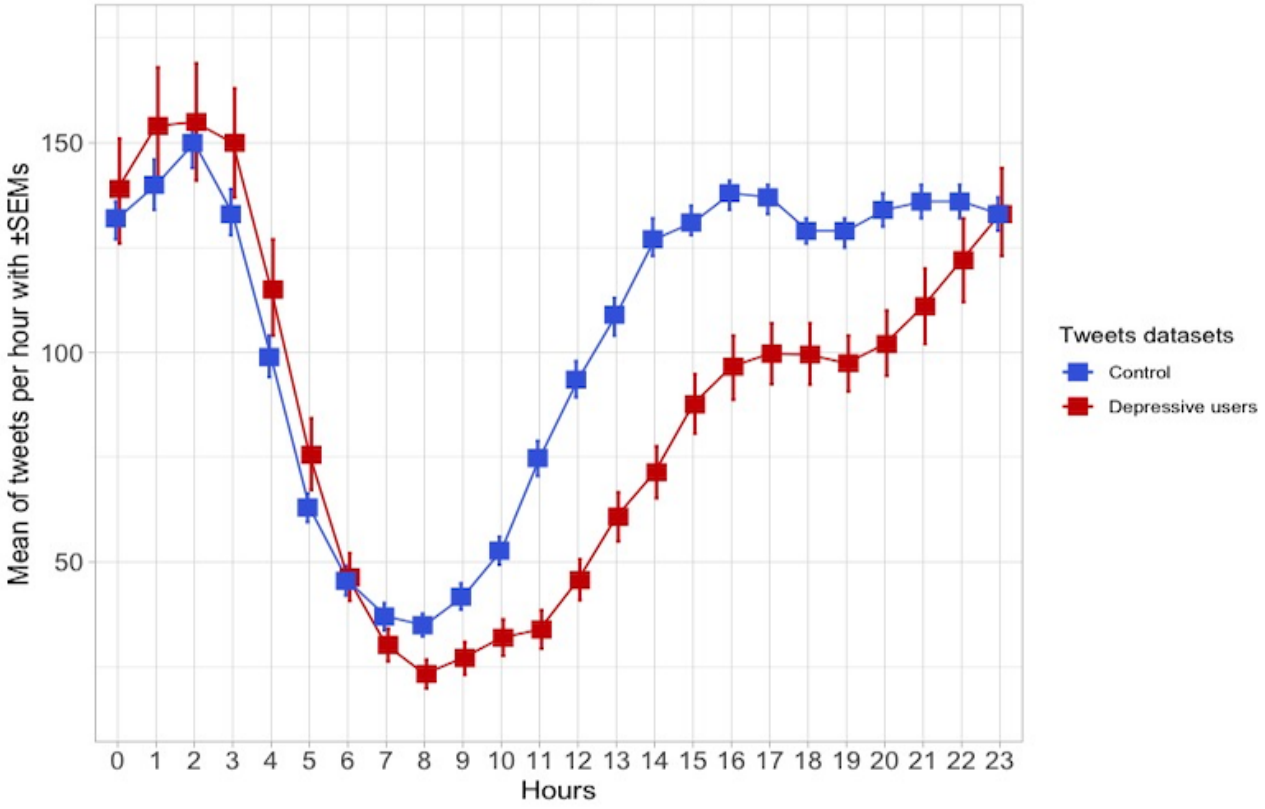
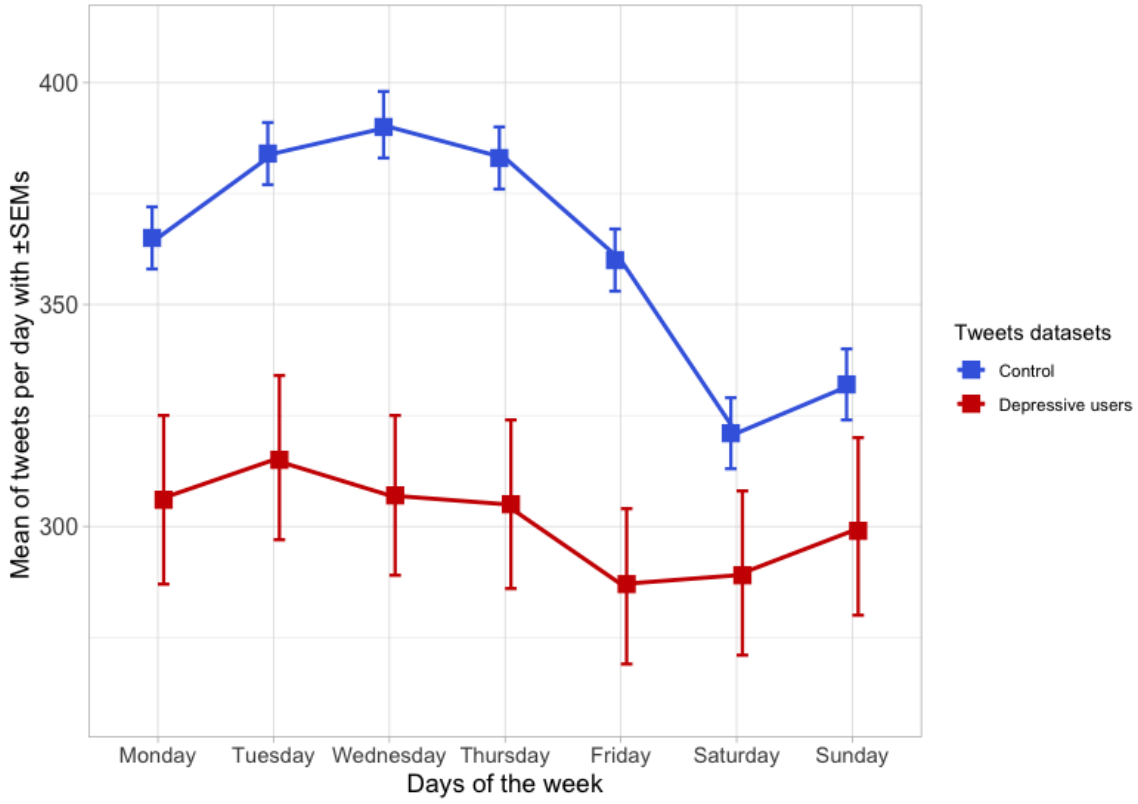


Figure 3. Number of tweets and retweets throughout the week of the control and depressive users datasets (mean±standard error of mean). SEM: standard error of mean.



XSL•FO
RenderX

Part-of-Speech

As for the analysis of POS corresponding to the number of words by grammatical categories in each tweet, we compared the 3 datasets of tweets: the control, depressive users, and depressive tweets datasets. As previously stated, the tweets of the depressive tweets dataset were removed from the depressive users dataset. The frequencies of words in each group are shown in Table 2. The number of nouns used in the control group almost doubles that of the depressive users dataset. By contrast, verbs are more frequently used in the depressive users dataset than in the control dataset. There were statistically significant differences between the control and the depressive users datasets (χ^2_7=1,242,600; *P*<.001), between the control and the depressive tweets datasets (χ^2_7=2,105.7; *P*<.001), and between the depressive users and the depressive tweets datasets (χ^2_7=15,888; *P*<.001).

In relation to the different types of pronouns in the control dataset, we detected 396,181 personal pronouns (51.38%; 396,181/770,955), the first-person singular (38.37%; 152,013/396,181) being the most used. A similar profile was observed in the depressive users dataset, where 124,614 pronouns were found (55.16%; 124,614/225,913), the first-person singular remaining the most used (57.59%; 71,768/124,614). In the depressive tweets dataset, 865 personal pronouns (53.16%; 865/1,627) were identified, and the first-person singular pronoun was by far the most used (80.00%; 692/865). The frequencies of personal pronouns in the different datasets are shown in Figure 4. There were statistically significant differences between the control and the depressive users datasets (χ^2_5=15,912; *P*<.001), between the control and the depressive tweets datasets (χ^2_5=638.7; *P*<.001), and between the depressive users and the depressive tweets datasets (χ^2_5=183.9; *P*<.001).

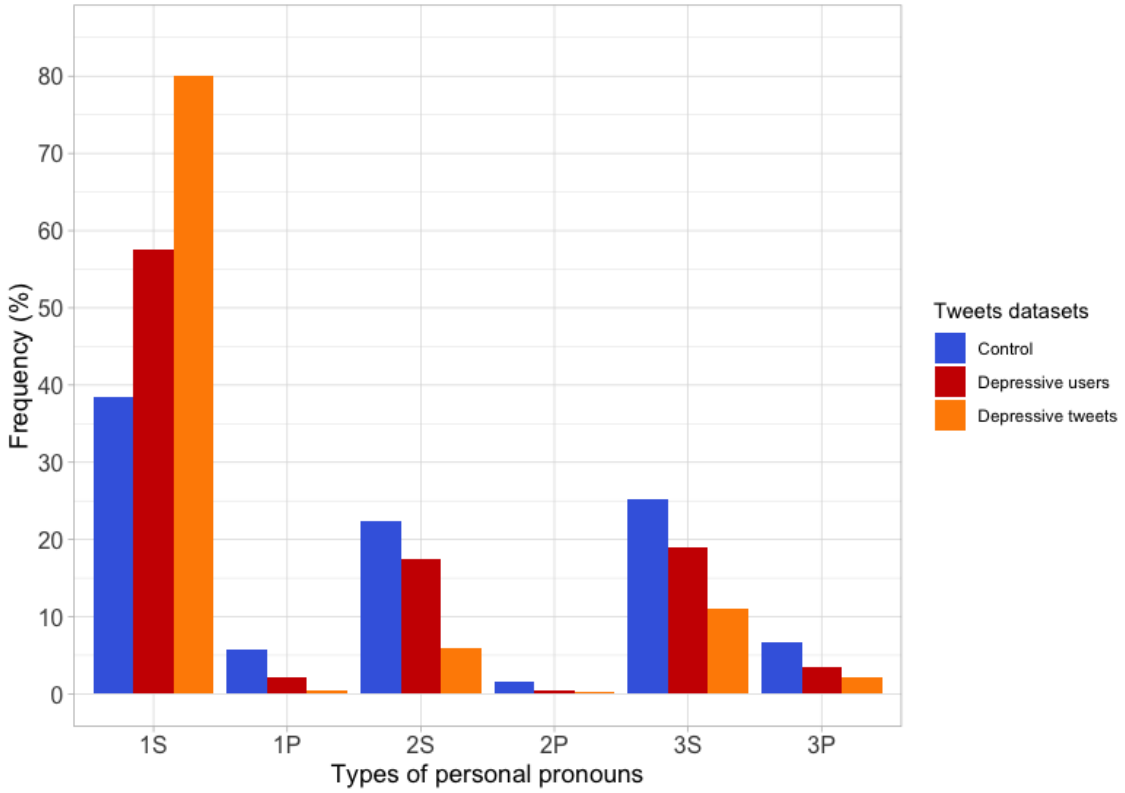
In relation to the number of characters per tweet, the mean of characters per tweet in the control and depressive users datasets was 83.48 (SD 40.57) and 65.76 (SD 36.99) characters, respectively, with statistically significant differences between them (t_{213770}=161.6; *P*<.001). On the other hand, the mean in the depressive tweets dataset was 67.51 (SD 38.28), which was not statistically significant and different in comparison with the depressive dataset ($t_{1012.3}$=1.45; *P*=.15). The 200 most frequent words that appeared in the control and depressive users datasets are depicted in the 2 word clouds shown in Multimedia Appendix 1. The 10 most frequent words that appeared in the control dataset were the following: *hoy* (today), *día* (day), *ver* (to see), *quiero* (I want), *gracias* (thank you), *mejor* (better), *siempre* (always), *vida* (life), *ahora* (now), and YouTube. In the depressive users dataset, the 10 most frequent words were the following: *quiero* (I want), *vida* (life), *siempre* (always), *siento* (I feel), *nadie* (nobody), *mierda* (shit), *never* (nunca), and *día* (day). It should be noted that in the depressive tweets dataset, although there are several words in common with the depressive users dataset, we can find additional words that are not present in the other datasets, such as *vacío/a* (empty), *matar* (to kill), *desaparecer* (to disappear), *suicidar* (commit suicide), *muerta* (dead), *desastre* (disaster), *inútil* (useless,), *deprimida* (depressed as state in women), *depresiva* (depressed as a condition in women), and *insomnio* (insomnia). The word cloud of the depressive tweets dataset is shown in Multimedia Appendix 2. In relation to the use of links, hashtags, and mentions in tweets, the frequency of them in the control and depressive users datasets were 35.32% (251,728/712,584), 13.13% (93,575/712,588), and 44.00% (313,574/712,577) and 18.07% (25,475/140,946), 1.44% (2030/140,946), and 9.27% (13,060/140,942), respectively. The number of tweets, including emojis, were 13.61% (97,038/712,589) in the control dataset and 5.72% (8069/140,947) in the depressive users dataset.

Table 2. Part-of-Speech (POS) frequencies in tweets of control, depressive users, and depressive tweets datasets.

Type of POS	POS in the control dataset, n (%)	POS in the depressive users dataset, n (%)	POS in the depressive tweets dataset, n (%)
Noun	2,298,544 (28.48)	270,104 (17.77)	1776 (15.07)
Verb	1,660,700 (20.58)	400,755 (26.36)	3391 (28.77)
Pronouns	770,955 (9.55)	225,913 (14.86)	1627 (13.80)
Adjectives	593,327 (7.35)	83,089 (5.47)	588 (4.99)
Determiner	1,068,130 (13.23)	177,795 (11.70)	1342 (11.39)
Adverbs	496,988 (6.16)	140,963 (9.27)	1351 (11.46)
Adpositions	854,573 (10.59)	123,867 (8.15)	1052 (8.93)
Conjunctions	327,852 (4.06)	97,541 (6.42)	659 (5.59)

XSL•FO
RenderX

Figure 4. Frequency of the different types of personal pronouns in the control, depressive users, and depressive tweets datasets. 1S: first-person singular; 1P: first-person plural; 2S: second-person singular; 2P: second-person plural; 3S: third-person singular; 3P: third-person plural.

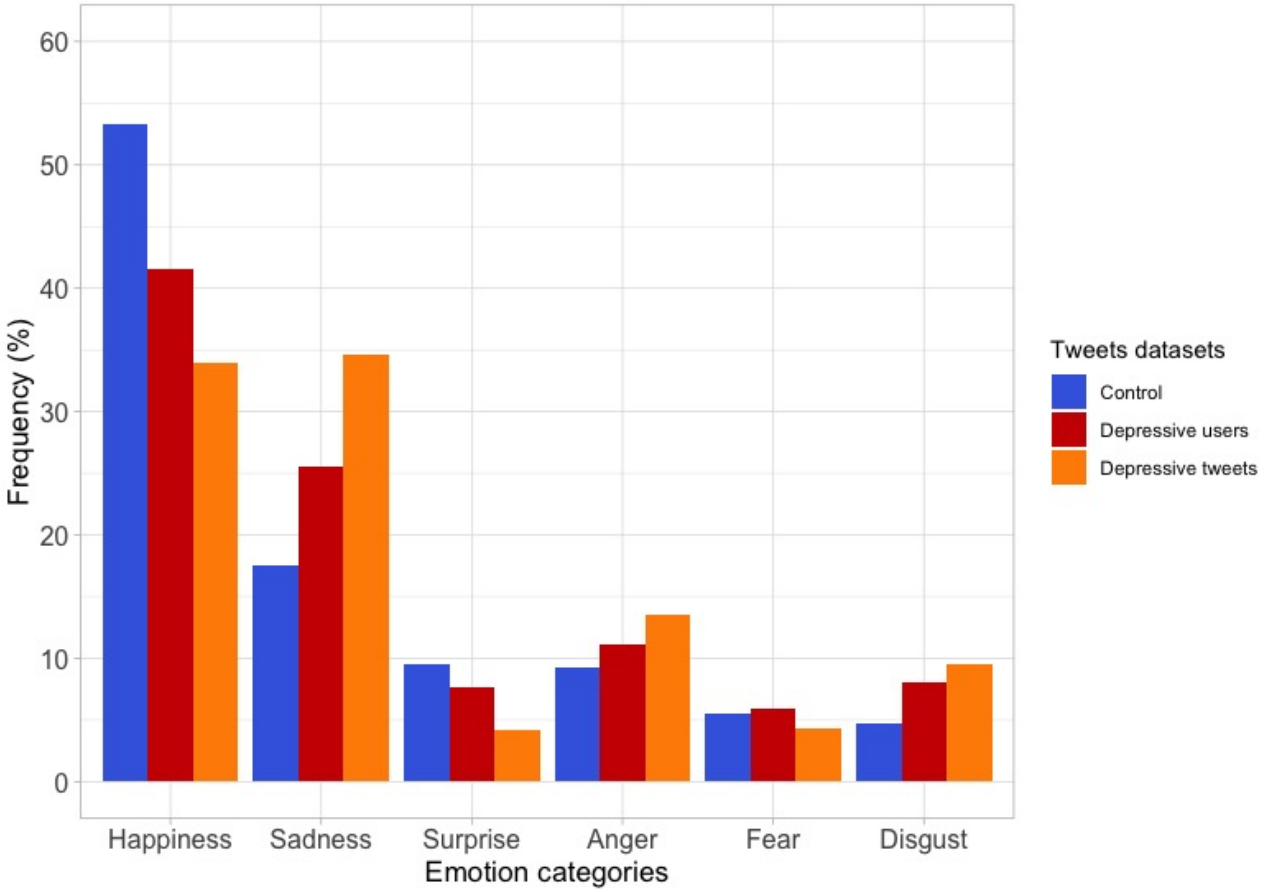


Emotion Analysis

Regarding the distribution of emotions, in the control dataset and in the depressive users dataset, the most frequent emotion was happiness (53.30%; 203,029/380,874 and 41.60%; 40,535/97,425) followed by sadness, which was more frequent in the depressive users dataset (17.59%; 67,033/380,874 and 25.49%; 24,834/97,425). In the depressive tweets dataset, the most frequent emotion was sadness (34.00%; 303/891). There were statistically significant differences between the control and the depressive users datasets (χ^2_5=6838.2; P<.001), between the control and the depressive tweets datasets (χ^2_5=296.8; P<.001), and between the depressive users and the depressive tweets datasets (χ^2_5=65.6; P<.001). The frequencies of the different emotions are shown in Figure 5.

XSL•FO
RenderX

Figure 5. Frequency distributions of emotion categories in the tweets of the 3 datasets.



Negation Words

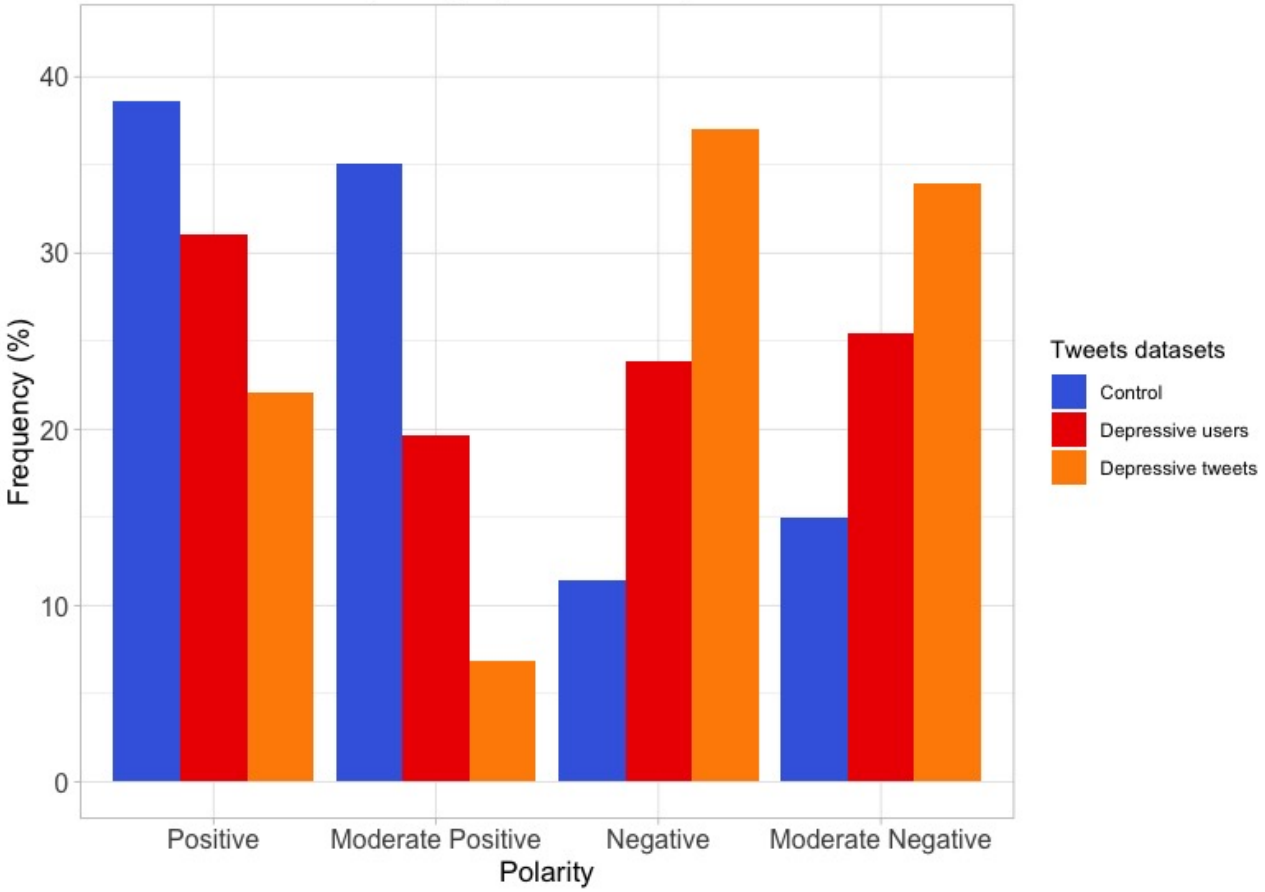
Regarding the use of negation words, they were detected in 21.74% (154,953/712,588) of the tweets in the control dataset, in 34.15% (48,137/140,946) of the depressive users dataset, and in 45.50% (455/1000) of the depressive tweets dataset. The mean of negation words was 0.28 (SD 0.59) in the control dataset, it was 0.49 (SD 0.82) in the depressive users dataset, and it was 0.67 (SD 0.91) in the depressive tweets dataset. There were statistically significant differences between the control and the depressive users datasets (Mann-Whitney U=4.3657e+10; P<.001), between the control and the depressive tweets datasets (Mann-Whitney U=266,990,000; P<.001), and between the depressive users and the depressive tweets datasets (Mann-Whitney U=62,002,000; P<.001).

Polarity Analysis

In relation to the polarity of tweets, 2 analyses were performed using 2 Spanish sentiment lexicons: the Senti Lexicon (including positive and negative categories) and the SentiCo Polarity (including positive, moderate positive, moderate negative, and negative categories). According to the Senti Lexicon, the analysis of tweets showed that the control dataset shows polarity in 33.47% (245,367/733,029) of the tweets, being positive in 56,54% (138,726/245,367) of them. In contrast, the depressive users dataset shows polarity in 41,31% (61,132/147,996) of the tweets, being positive in 46.14% (28,205/61,132) of them. Finally, the depressive tweets dataset shows polarity in 58.90% (589/1000) of the tweets, with positive polarity in 34.97% (206/589) of them. There were statistically significant differences between the control and the depressive users datasets (χ^2_1=2134; P<.001), between the control and the depressive tweets datasets (χ^2_1=110.3; P<.001), and between the depressive users and the depressive tweets datasets (χ^2_1=28.8; P<.001). When using the SentiCo Polarity tool, the control dataset presented 20.97% (152,228/725,717) of tweets with polarity, 29.32% (42,820/146,033) in the depressive users and 33.34% (348/1,044) in the depressive tweets dataset. The distributions of polarities are shown in Figure 6. There were statistically significant differences between the control and the depressive users datasets (χ^2_3=8820.8, P<.001), between the control and the depressive tweets datasets (χ^2_3=308.8; P<.001), and between the depressive users and the depressive tweets datasets (χ^2_3=52.4; P<.001).

XSL•FO
RenderX

Figure 6. Polarities of the tweets according to the SentiCo Polarity tool in the 3 datasets.



Discussion

Principal Findings

The diagnosis of depression is complex because of the heterogeneous nature of this disease and the diverse manifestation of the symptoms among individuals, which result in a great number of depressive disorder cases that are undetected and untreated, making the prevention, diagnosis, and treatment of the depressive disorders a complicated task [15,49,50]. For these reasons and taking into account that people diagnosed with depression are increasing worldwide, new strategies for detecting and monitoring this disease would be very useful. In this study, we analyzed the behavioral and linguistic patterns of tweets in Spanish that suggest signs of depression. The results contribute to the growing body of scientific literature that analyzes the messages posted on social media using languages other than English. We have introduced a new approach that comprises analyzing the timelines of self-qualified depressed users, as well as their tweets related to depression, which are manually selected. Our results show that the tweets of depressive users have different features in comparison with those of a control dataset, even when their tweets that are not related to depression are considered (depressive users dataset). In addition, the differences with the control dataset become more evident when we consider the manual selection of tweets related to depression (depressive tweets dataset).

Different Distributions of Tweets Over Time

As for the distribution of tweets over time, the users of the depressive dataset, although being less active in using Twitter, used to be more active during the night than the users of the control dataset. This can be explained as a result of insomnia, one of the most frequent symptoms of depression. This finding is consistent with previous studies carried out with English speakers, which demonstrated that individuals with depression are more active during the night [20]. Moreover, the daily mood changes, such as the morning and evening worsening that are typical in several forms of depression, could explain the lower activity of the depressive users [51]. In relation to the distribution of tweets throughout the week, the users of the depressive dataset showed a more regular activity throughout the week, tending to be more active on Saturdays, Sundays, and Mondays than those of the control dataset, whose activity showed a drop during the weekend. This trend may be related to the lowered social activity of the people suffering from depressive disorders, having a reduced participation in social leisure activities during the weekend and spending more time at home, sharing their feelings and thoughts on social media platforms [16].

Different Style of Writing

The analysis of POS and the number of words by grammatical categories show that, generally, the users of the depressive dataset used more verbs, adverbs, and pronouns but less nouns than the control dataset. The same features are also present in the depressive tweets dataset. These findings suggest that the

XSL•FO
RenderX

language of people suffering from depression is characterized by a different style of writing that some authors describe as poorly structured, indicating less interest in what surrounds them, people, objects, or things [52]. They focus on talking about actions, and this is correlated with sensitive disclosure. Consistent with many previous studies [20,30,35,53-55], the use of first-person singular pronoun is more frequent among the users of the depressive dataset, with respect to those of the control dataset, and this difference increases in the depressive tweets dataset. The increased use of this pronoun demonstrates the attention to self-focus that is associated with the negative emotional states of depression and the reduced attentional resources, highlighting the psychological distancing to connect with others [56]. This social isolation may also explain that the first- and second-person plural pronouns are the least used. Language can be used as a measure of different individual features, on the basis of the fact that people's word choice is stable over time and consistent across topics or context. For this reason, the language style appears to be a useful predictor of some mental health conditions, such as depression [20,35]. In addition, the number of characters written in the depressive users and depressive tweets datasets was smaller than the number of characters written in the control dataset, and this might be related to reduced interest and poorer language. According to the most frequent words that appeared in the depressive users and depressive tweets datasets, there are specific words that are linked to clinical symptoms and the way that depressive patients word their mood, such as words that may be related to suicide ideation. Consequently, they can be used as a signal to detect potentially depressed users on Twitter [36]. Similarly, we observed the frequent use of adjectives in feminine form in the depressive tweets dataset, which would suggest that a high proportion of the depressive users are women, a fact that is in agreement with clinical and epidemiological evidences [8,11,12,42].

Predominant Emotions

Emotions are one of the key aspects that characterize many mental health conditions and, particularly, when people are suffering from depression. An analysis of the 6 emotions that are commonly considered (happiness, sadness, surprise, anger, fear, and disgust) [47] was performed to determine the existence of differences among the datasets. Happiness is the most frequent emotion in the control and depressive users datasets, although an important reduction was observed in the depressive tweets dataset. The surprise emotion is less frequent in depressive users and, specially, in the depressed tweets datasets than the control dataset, and this fact can be related to the depressive mood, in which there is a decrease in interest in almost everything. Fear does not seem to be a differential emotion in the groups of tweets analyzed in this study.

Regarding negative emotions, we observed an increase in the frequency of words related to the sadness emotion in the depressive tweets dataset, doubling that of the control dataset. This feature had also been observed in other studies [14,35,57]. Moreover, anger is more frequent in the depressive user and depressive tweets datasets than in the control dataset. Although Twitter is used many times for expressing anger about personal or political aspects, this emotion is particularly frequent in patients suffering from depression, who tend to feel irritable, wronged, or angry at the world [14,16,35,58]. At the same time, disgust, an emotion that is known to be associated with the depressive disorders [59], was found to be more frequent in the depressive users and depressive tweets datasets.

Negative Focused Emotion Language

In our analysis, the presence of negation words is more frequent in the depressed users (34.15%; 48,137/140,946) and depressive tweets (45.50%; 455/1000) datasets than in the control dataset (21.74%; 154,953/712,588), indicating that there is an increased use of negatively focused emotion language, which is typical in depressive patients and feelings [31,54,55,60].

Negative Polarity

The classification of tweets, on the basis of the emotional positivity or negativity of their words, is another analysis that has been carried out. In this study, we used 2 types of polarity lexicons, the Senti Lexicon (SentiLex) and the Sentico Polarity (SentiCo). In both cases, the negative polarity was higher in the depressive users and depressive tweets datasets, even tripling the negativeness of the control dataset when using the Sentico Polarity lexicon. These findings are consistent with other studies, indicating that people suffering from depression tend to focus more on negative aspects of their life [20,35], and thus their tweets contain much more negative emotional words compared with the control dataset [14]. In addition, the self-focus state that characterizes depression is associated with negative emotions [32,56,57].

Limitations and Future Directions

This study presents some limitations that have to be pointed out. On the one hand, the tweets of the depressive datasets come only from Twitter users who speak publicly about feelings and emotions that can be related with depression. This is an indirect and inaccurate way of detecting users suffering from depressive disorders. Without clinically assessing these people, there is no way to verify if the diagnosis is genuine or if they suffer from another mental disorder. On the other hand, it is possible that Twitter users self-disclose their mental health using words or expressions not included in the list of keywords used in this study for streaming tweets about depression [22,61-63]. In this respect, it is possible that a wider list could have yielded a greater coverage [21,36]. Privacy policies of social media restrict the access to users who did not grant access to their profile, and this may have generated biases in the composition of the depressive users and the depressive tweets datasets. In addition, tweets may incorporate biases because of the self-management and anonymity of the Web-based identities [61]. Moreover, Twitter users may be not be representative of the general population, and some studies have shown that they are often urban people with high levels of education [64-66]. More information about the socioeconomic and demographic details of Twitter users is needed [67]. The control dataset was a randomly selected sample of Twitter users, and it is consequently representative of the users of this social media. However, there is a possibility that users in this group may also have depression or other mental illness even though they did not mention this in their profile description. There is also the

XSL•FO
RenderX

possibility that the users included in the control group are fake accounts. Only original tweets were analyzed, and perhaps retweets, which are not included in our linguistic study, reflect users' emotions that can be related to depression status [68]. Finally, depression is a very complex mental disorder, and our study only provides a general observation of this disorder. Additional research might be carried out to examine specific depression types and determine if there are social media features that can contribute to classifying users or tweets to the different diagnosis of depression [69]. Similarly, in future works, we plan to study the linguistic features and the behavioral patterns of depression in different linguistics contexts. The possible relationship between depression and seasonality could be of interest for future studies in the context of monitoring Twitter activity [70].

Conclusions

The prevalence of common mental disorders worldwide, such as depression, requires the ability of health care systems to provide adequate diagnosis, monitoring, and treatment. The wide popularity of social media platforms introduces new opportunities for the screening of depression. The introduction of new methods of analysis for the automatic detection of signals of depression on social media platforms, such as Twitter or Facebook, has the potential of being used as a complementary tool for the assessment of these patients, assisting health care professionals in the detection and monitoring of mental health disorders. Although the analysis of tweets as a way to determine the existence of depression cannot be used as a replacement for diagnosis, it has the potential as a screening tool for depressive disorders, with a lower cost than other traditional procedures. In addition, it can be helpful to health professionals for managing and monitoring patients more efficiently. Similarly, it can be useful for particular patients, as they feel more comfortable disclosing their symptoms on Twitter than in clinical settings. In this study, we have shown that several behavioral and linguistic features of the tweets in Spanish can be used as a complementary tool to detect signals of depression of their authors, corroborating and extending the findings obtained by studies carried out on English tweets. As we described in this study, signs of depression of Twitter users are not exclusively spotted by identifying and analyzing tweets that explicitly mention expressions related to depression. Moreover, Twitter users who are potentially suffering from depression globally modify the core traits of their language, independently from the fact that the tweets are related or not related to the expression of depression. On the basis of these changes, these users can be monitored and supported. The results of this paper, jointly with other studies on the matter, support the potential of social media as an important instrument for extending and enhancing mental health services available to people with mental disorders. By means of interdisciplinary collaborations, it is possible to develop digital apps and services providing personalized alerts and psychosocial support in the mental health domain.

Acknowledgments

We received support from the Agency for Management of University and Research Grants in Catalonia (Spain) for the incorporation of new research personnel (FI2016), the European Union H2020 Research and Innovation Programme 2014-2020 under grant agreement number 634143 (MedBioinformatics: Creating medically driven integrative bioinformatics applications focused on oncology and central nervous system disorders and their comorbidities). The Research Programme on Biomedical Informatics is a member of the Spanish National Bioinformatics Institute, PRB2-ISCIII, and it is supported by grant PT17/0009/0014.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Word clouds showing the 200 most frequent words in the control (left) and depressive users (right) datasets.

[PNG File, 396KB - jmir_v21i6e14199_app1.png]

Multimedia Appendix 2

Word cloud showing the 200 most frequent words in the depressive tweets dataset.

[PNG File, 166KB - jmir_v21i6e14199_app2.png]

References

1. World Health Organization. 2013. Mental Health Action Plan 2013-2020 URL: http://apps.who.int/iris/bitstream/10665/89966/1/9789241506021_eng.pdf?ua=1 [accessed 2018-04-25] [WebCite Cache ID 778TsWziL]
2. Marcus M, Yasamy MT, van Ommeren M, Chisholm D, Saxena S, WHO Department of Mental Health and Substance Abuse. World Health Organization. 2012. Depression: A Global Public Health Concern URL: http://www.who.int/mental_health/management/depression/who_paper_depression_wfmh_2012.pdf [accessed 2018-04-25] [WebCite Cache ID 778NmzRq1]

XSL•FO
RenderX

3. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the global burden of disease study 2013. Lancet 2015 Aug 22;386(9995):743-800 [FREE Full text] [doi: 10.1016/S0140-6736(15)60692-4] [Medline: 26063472]
4. Trautmann S, Rehm J, Wittchen HU. The economic costs of mental disorders: do our societies react appropriately to the burden of mental disorders? EMBO Rep 2016;17(9):1245-1249 [FREE Full text] [doi: 10.15252/embr.201642951] [Medline: 27491723]
5. Patel V, Chisholm D, Parikh R, Charlson FJ, Degenhardt L, Dua T, et al. Global priorities for addressing the burden of mental, neurological,substance use disorders. In: Patel V, Chisholm D, Dua T, Laxminarayan R, Medina-Mora ME, Vos T, editors. Disease Control Priorities: Mental, Neurological, and Substance Use Disorders, Third Edition (Volume 4). Washington, DC: The World Bank; 2016:1-27.
6. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the global burden of disease study 2010. Lancet 2013 Nov 9;382(9904):1575-1586. [doi: 10.1016/S0140-6736(13)61611-6] [Medline: 23993280]
7. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J Med Internet Res 2017 Dec 29;19(6):e228 [FREE Full text] [doi: 10.2196/jmir.7215] [Medline: 28663166]
8. World Health Organization. 2018. Depression: Key Facts URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2018-04-23] [WebCite Cache ID 778TC0BMP]
9. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 2006 Nov;3(11):e442 [FREE Full text] [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]
10. Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. Lancet Psychiatry 2016 Feb;3(2):171-178. [doi: 10.1016/S2215-0366(15)00505-2] [Medline: 26851330]
11. World Health Organization. 2017. Depression and Other Common Mental Disorders: Global Health Estimates URL: http://apps.who.int/iris/bitstream/10665/254610/1/WHO-MSD-MER-2017.2-eng.pdf?ua=1 [accessed 2018-04-25] [WebCite Cache ID 77CsSorTr]
12. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med 2013 Nov;10(11):e1001547 [FREE Full text] [doi: 10.1371/journal.pmed.1001547] [Medline: 24223526]
13. Paul MJ, Dredze M. You Are What You Tweet: Analyzing Twitter for Public Health. In: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media. 2011 Presented at: AAAI'11; July 17-21, 2011; Barcelona, Spain p. 265-272.
14. Park M, Cha C, Cha M. Depressive Moods of Users Portrayed in Twitter. In: Proceedings of the ACM SIGKDD Workshop on Health Informatics. 2012 Presented at: HI-KDD'12; August 12-16, 2012; Beijing, China p. 1-8.
15. Nguyen T, O'Dea B, Larsen M, Phung D, Venkatesh S, Christensen H. Using linguistic and topic analysis to classify sub-groups of online depression communities. Multimed Tools Appl 2015 Dec 21;76(8):10653-10676. [doi: 10.1007/s11042-015-3128-x]
16. Cavazos-Rehg PA, Krauss MJ, Sowles S, Connolly S, Rosas C, Bharadwaj M, et al. A content analysis of depression-related Tweets. Comput Human Behav 2016;54:351-357 [FREE Full text] [doi: 10.1016/j.chb.2015.08.023] [Medline: 26392678]
17. Coppersmith G, Dredze M, Harman C, Hollingshead K. From ADHD to SAD: Analyzing the Language of Mental Health on Twitter Through Self-Reported Diagnoses. In: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2015 Presented at: CLPSYCH'15; June 5, 2015; Denver, Colorado p. 1-10 URL: https://www.aclweb.org/anthology/W15-1201 [doi: 10.3115/v1/W15-1201]
18. Naslund JA, Aschbrenner KA, McHugo GJ, Unützer J, Marsch LA, Bartels SJ. Exploring opportunities to support mental health care using social media: a survey of social media users with mental illness. Early Interv Psychiatry 2019 Jun;13(3):405-413. [doi: 10.1111/eip.12496] [Medline: 29052947]
19. Berry N, Lobban F, Belousov M, Emsley R, Nenadic G, Bucci S. #WhyWeTweetMH: understanding why people use Twitter to discuss mental health problems. J Med Internet Res 2017 Dec 5;19(4):e107 [FREE Full text] [doi: 10.2196/jmir.6173] [Medline: 28381392]
20. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting Depression via Social Media. In: Proceedings of the Seventh International Conference on Weblogs and Social Media. 2013 Presented at: AAAI'13; July 8-11, 2013; Cambridge, MA p. 128-138.
21. Wilson ML, Ali S, Valstar MF. Finding Information About Mental Health in Microblogging Platforms: A Case Study of Depression. In: Proceedings of the 5th Information Interaction in Context Symposium. 2014 Presented at: IIiX'14; August 26-30, 2014; Regensburg, Germany p. 8-17. [doi: 10.1145/2637002.2637006]
22. Coppersmith G, Dredze M, Harman C. Quantifying Mental Health Signals in Twitter. In: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality. 2014 Presented at: CLPSYCH'14; June 5, 2014; Baltimore, Maryland p. 51-60. [doi: 10.3115/v1/W14-3207]

XSL•FO
RenderX

23. Radzikowski J, Stefanidis A, Jacobsen KH, Croitoru A, Crooks A, Delamater PL. The measles vaccination narrative in Twitter: a quantitative analysis. JMIR Public Health Surveill 2016;2(1):e1 [FREE Full text] [doi: 10.2196/publichealth.5059] [Medline: 27227144]
24. Stefanidis A, Vraga E, Lamprianidis G, Radzikowski J, Delamater PL, Jacobsen KH, et al. Zika in Twitter: temporal variations of locations, actors, and concepts. JMIR Public Health Surveill 2017 Apr 20;3(2):e22 [FREE Full text] [doi: 10.2196/publichealth.6925] [Medline: 28428164]
25. Finch KC, Snook KR, Duke CH, Fu K, Tse ZT, Adhikari A, et al. Public health implications of social media use during natural disasters, environmental disasters, and other environmental concerns. Nat Hazards 2016 Apr 19;83(1):729-760. [doi: 10.1007/s11069-016-2327-8]
26. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
27. Conway M, O'Connor D. Social media, big data, and mental health: current advances and ethical implications. Curr Opin Psychol 2016 Jun;9:77-82 [FREE Full text] [doi: 10.1016/j.copsyc.2016.01.004] [Medline: 27042689]
28. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
29. Statista. 2018. Number of Monthly Active Twitter Users Worldwide From 1st Quarter 2010 to 2nd Quarter 2018 (in Millions) URL: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ [accessed 2018-10-20] [WebCite Cache ID 778OhlXHI]
30. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 2009 Dec 8;29(1):24-54 [FREE Full text] [doi: 10.1177/0261927X09351676]
31. Ramirez-Esparza N, Chung CK, Kacewicz E, Pennebaker JW. The Psychology of Word Use in Depression Forums in Englishin Spanish: Texting Two Text Analytic Approaches. In: Proceedings of the 2nd International Conference on Weblogs and Social Media. 2008 Presented at: AAAI'08; March 30–April 2, 2008; Seattle, Washington URL: https://www.aaai.org/Papers/ICWSM/2008/ICWSM08-020.pdf
32. Chung C, Pennebaker J. The psychological functions of function words. In: Fiedler K, editor. Social Communication: Frontiers of Social Psychology. First Edition. New York: Psychology Press; 2007:343-359.
33. 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]
34. 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 Oct 29;13:508 [FREE Full text] [doi: 10.1186/1471-2407-13-508] [Medline: 24168075]
35. Pavalanathan U, De Choudhury M. Identity Management and Mental Health Discourse in Social Media. In: Proceedings of the 24th International Conference on World Wide Web. 2015 Presented at: WWW'15 Companion; May 18-22, 2015; Florence, Italy p. 315-321. [doi: 10.1145/2740908.2743049]
36. De Choundhury M, Counts S, Horvitz E. Social Media as a Measurement Tool of Depression in Populations. In: Proceedings of the 5th Annual ACM Web Science Conference. 2013 Presented at: WebSci'13; May 2-4, 2013; Paris, France p. 47-56 URL: https://dl.acm.org/citation.cfm?id=2464480 [doi: 10.1145/2464464.2464480]
37. Coppersmith G, Harman C, Dredze M. Measuring Post Traumatic Stress Disorder in Twitter. In: Proceedings of the 8th International Conference on Weblogs and Social Media. 2014 Presented at: AAAI'14; June 1-4, 2014; Ann Arbor, Michigan p. 579-582 URL: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewFile/8079/8082
38. Birnbaum ML, Ernala SK, Rizvi AF, De Choudhury M, Kane JM. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. J Med Internet Res 2017 Dec 14;19(8):e289 [FREE Full text] [doi: 10.2196/jmir.7956] [Medline: 28807891]
39. Arseniev-Koehler A, Lee H, McCormick T, Moreno MA. #Proana: pro-eating disorder socialization on Twitter. J Adolesc Health 2016;58(6):659-664. [doi: 10.1016/j.jadohealth.2016.02.012] [Medline: 27080731]
40. Statista. 2019. Leading Countries Based on Number of Twitter Users as of January 2019 (in Millions) URL: https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ [WebCite Cache ID 77Ctc6VPr]
41. Twitter Developer. URL: https://developer.twitter.com/en.html [accessed 2018-11-10] [WebCite Cache ID 778Qhq4Gu]
42. American Psychiatric Association. Diagnostic And Statistical Manual Of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Publishing; 2013.
43. Natural Language Tool Kit. URL: https://www.nltk.org/api/nltk.tokenize.html [accessed 2018-10-10] [WebCite Cache ID 778Rzo6pe]
44. Padró L, Stanilovsky E. FreeLing 3.0: Towards Wider Multilinguality. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation. 2012 Presented at: LREC'12; May 21-27, 2012; Istanbul, Turkey p. 2473-2479 URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/430_Paper.pdf
45. Perez-Rosas V, Banea C, Mihalcea R. Learning Sentiment Lexicons in Spanish. In: Proceedings of the Eighth International Conference on Language Resources and Evaluation. 2012 Presented at: LREC'12; May 21-27, 2012; Istanbul, Turkey p. 3077-3081 URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/1081_Paper.pdf

XSL•FO
RenderX

46. Cruz FL, Troyano JA, Pontes B, Ortega FJ. Building layered, multilingual sentiment lexicons at synset and lemma levels. Expert Syst Appl 2014 Oct;41(13):5984-5994 [FREE Full text] [doi: 10.1016/j.eswa.2014.04.005]
47. Ekman P, Friesen WV, O'Sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, et al. Universals and cultural differences in the judgments of facial expressions of emotion. J Pers Soc Psychol 1987 Oct;53(4):712-717. [doi: 10.1037/0022-3514.53.4.712] [Medline: 3681648]
48. Sidorov G, Miranda-Jiménez S, Viveros-Jiménez F, Gelbukh A, Castro-Sánchez NA, Castillo F, et al. Empirical Study of Opinion Mining in Spanish Tweets. In: Proceedings of the 11th Mexican International Conference on Artificial Intelligence. 2012 Presented at: MICAI'12; October 27-November 4, 2012; San Luis Potosí, Mexico p. 1-4.
49. Nambisan P, Luo Z, Kapoor A, Patrick TB, Cisler RA. Social Media, Big Data, and Public Health Informatics: Ruminating Behavior of Depression Revealed Through Twitter. In: Proceedings of the 2015 48th Hawaii International Conference on System Sciences. 2015 Presented at: HICSS'15; January 5-8, 2015; Kauai, Hawaii p. 2906-2913. [doi: 10.1109/HICSS.2015.351]
50. Cassano P, Fava M. Depression and public health: an overview. J Psychosom Res 2002 Oct;53(4):849-857. [Medline: 12377293]
51. Morris DW, Rush AJ, Jain S, Fava M, Wisniewski SR, Balasubramani GK, et al. Diurnal mood variation in outpatients with major depressive disorder: implications for DSM-V from an analysis of the sequenced treatment alternatives to relieve depression study data. J Clin Psychiatry 2007 Sep;68(9):1339-1347. [doi: 10.4088/JCP.v68n0903] [Medline: 17915971]
52. De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M. Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16; May 7-12, 2016; San Jose, California p. 2098-2110 URL: http://europepmc.org/abstract/MED/29082385 [doi: 10.1145/2858036.2858207]
53. Bucci W, Freedman N. The language of depression. Bull Menninger Clin 1981 Jul;45(4):334-358. [Medline: 6176285]
54. Rude S, Gortner EM, Pennebaker J. Language use of depressed and depression-vulnerable college students. Cogn Emot 2004;18(8):1121-1133. [doi: 10.1080/02699930441000030]
55. Pyszczynski T, Greenberg J. Self-regulatory perseveration and the depressive self-focusing style: a self-awareness theory of reactive depression. Psychol Bull 1987 Jul;102(1):122-138. [doi: 10.1037/0033-2909.102.1.122] [Medline: 3615702]
56. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language use: our words, our selves. Annu Rev Psychol 2003;54:547-577. [doi: 10.1146/annurev.psych.54.101601.145041] [Medline: 12185209]
57. Sonnenschein AR, Hofmann SG, Ziegelmayer T, Lutz WL. Linguistic analysis of patients with mood and anxiety disorders during cognitive behavioral therapy. Cogn Behav Ther 2018;47(4):315-327. [doi: 10.1080/16506073.2017.1419505] [Medline: 29345528]
58. Painuly N, Sharan P, Mattoo SK. Relationship of anger and anger attacks with depression: a brief review. Eur Arch Psychiatry Clin Neurosci 2005 Aug;255(4):215-222. [doi: 10.1007/s00406-004-0539-5] [Medline: 16133740]
59. Neacsiu AD, Rompogren J, Eberle JW, McMahon K. Changes in problematic anger, shame, and disgust in anxious and depressed adults undergoing treatment for emotion dysregulation. Behav Ther 2018;49(3):344-359 [FREE Full text] [doi: 10.1016/j.beth.2017.10.004] [Medline: 29704965]
60. Bernard JD, Baddeley JL, Rodriguez BF, Burke PA. Depression, language, and affect: an examination of the influence of baseline depression and affect induction on language. J Lang Soc Psychol 2016;35(3):317-326. [doi: 10.1177/0261927X15589186]
61. De Choudhury M, Sharma SS, Logar T, Eekhout W, Nielsen RC. Gender and Cross-Cultural Differences in Social Media Disclosures of Mental Illness. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017 Presented at: CSCW'17; February 25-March 1, 2017; Portland, Oregon p. 353-369. [doi: 10.1145/2998181.2998220]
62. De Choudhury M, De S. Mental Health Discourse on Reddit: Self-Disclosure, Social Support, and Anonymity. In: Proceedings of the 8th International Conference on Weblogs and Social Media. 2014 Presented at: AAAI'14; June 1-4, 2014; Ann Arbor, Michigan p. 71-80 URL: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8075/8107
63. Mowery D, Smith H, Cheney T, Stoddard G, Coppersmith G, Bryan C, et al. Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study. J Med Internet Res 2017 Dec 28;19(2):e48 [FREE Full text] [doi: 10.2196/jmir.6895] [Medline: 28246066]
64. Eichstaedt JC, Schwartz HA, Kern ML, Park G, Labarthe DR, Merchant RM, et al. Psychological language on Twitter predicts county-level heart disease mortality. Psychol Sci 2015 Feb;26(2):159-169 [FREE Full text] [doi: 10.1177/0956797614557867] [Medline: 25605707]
65. Yom-Tov E, Johansson-Cox I, Lampos V, Hayward AC. Estimating the secondary attack rate and serial interval of influenza-like illnesses using social media. Influenza Other Respir Viruses 2015 Jul;9(4):191-199 [FREE Full text] [doi: 10.1111/irv.12321] [Medline: 25962320]
66. Paul MJ, Dredze M. Discovering health topics in social media using topic models. PLoS One 2014;9(8):e103408 [FREE Full text] [doi: 10.1371/journal.pone.0103408] [Medline: 25084530]
67. Chary M, Genes N, McKenzie A, Manini AF. Leveraging social networks for toxicovigilance. J Med Toxicol 2013 Jun;9(2):184-191 [FREE Full text] [doi: 10.1007/s13181-013-0299-6] [Medline: 23619711]

XSL•FO
RenderX

68. Seabrook EM, Kern ML, Fulcher BD, Rickard NS. Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. J Med Internet Res 2018 Dec 8;20(5):e168 [FREE Full text] [doi: 10.2196/jmir.9267] [Medline: 29739736]
69. Reece AG, Reagan AJ, Lix KL, Dodds PS, Danforth CM, Langer EJ. Forecasting the onset and course of mental illness with Twitter data. Sci Rep 2017 Dec 11;7(1):13006 [FREE Full text] [doi: 10.1038/s41598-017-12961-9] [Medline: 29021528]
70. Alvarez-Mon MA, Asunsolo Del Barco A, Lahera G, Quintero J, Ferre F, Pereira-Sanchez V, et al. Increasing interest of mass communication media and the general public in the distribution of Tweets about mental disorders: observational study. J Med Internet Res 2018 Dec 28;20(5):e205 [FREE Full text] [doi: 10.2196/jmir.9582] [Medline: 29807880]

Abbreviations

API: Application Programming Interface
POS: Part-of-Speech
WHO: World Health Organization

Edited by G Eysenbach; submitted 29.03.19; peer-reviewed by JL Oliveira, E Gabarron, X Chen, V Pereira-Sanchez; comments to author 23.04.19; revised version received 24.05.19; accepted 24.05.19; published 27.06.19.

Please cite as:
Leis A, Ronzano F, Mayer MA, Furlong LI, Sanz F
Detecting Signs of Depression in Tweets in Spanish: Behavioral and Linguistic Analysis
J Med Internet Res 2019;21(6):e14199
URL: http://www.jmir.org/2019/6/e14199/
doi:10.2196/14199
PMID:31250832

©Angela Leis, Francesco Ronzano, Miguel A Mayer, Laura I Furlong, Ferran Sanz. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Provision of Paid Web-Based Medical Consultation in China: Cross-Sectional Analysis of Data From a Medical Consultation Website

Yumei Li[1], MA; Xiangbin Yan[1,2], PhD; Xiaolong Song[3], PhD

[1]Harbin Institute of Technology, Harbin, China

[2]University of Science and Technology Beijing, Beijing, China

[3]School of Management Science and Engineering, Dongbei University of Finance and Economics, Dalian, China

Corresponding Author:
Xiangbin Yan, PhD
Harbin Institute of Technology
92, Xidazhi Street, Nangang District
Harbin,
China
Phone: 86 13351288123
Email: xbyan@hit.edu.cn

Abstract

Background: Web-based medical consultation, which has been adopted by patients in many countries, requires a large number of doctors to provide services. However, no study has provided an overall picture of the doctors who provide online services.

Objective: This study sought to examine doctors' participation in medical consultation practice in an online consultation platform. This paper aimed to address the following questions: (1) which doctors provide Web-based consultation services, (2) how many patients do the doctors get online, and (3) what price do they charge. We further explored the development of market segments in various departments and various provinces.

Methods: This study explored the dataset including all doctors providing consultation services in their spare time on a Chinese Web-based consultation platform. We also brought in statistics for doctors providing offline consultations in China. We made use of Bonferroni multiple comparison procedures and z test to compare doctors in each group.

Results: There are 88,308 doctors providing Web-based consultation in their spare time on Haodf, accounting for 5.25% (88,308/1,680,062) of all doctors in China as of September 23, 2017. Of these online doctors, 49.9% (44,066/88,308) are high-quality doctors having a title of chief physician or associate chief physician, and 84.8% (74,899/88,308) come from the top, level 3, hospitals. Online doctors had an average workload of 0.38 patients per doctor per day, with an online and offline ratio of 1:14. The average price of online consultation is ¥32.3. The online ratios for the cancer, internal medicine, ophthalmology-otorhinolaryngology, psychiatry, gynecology-obstetrics-pediatrics, dermatology, surgery, and traditional Chinese medicine departments are 16.1% (2,983/18,481), 4.4% (16,231/372,974), 6.3% (8,389/132,725), 9.5% (1,600/16,801), 5.8% (12,955/225,128), 18.0% (3,334/18,481), 10.8% (24,030/223,448), and 3.8% (8,393/22,3448), respectively. Most provinces located in eastern China have more than 4000 doctors online. The online workloads for doctors from most provinces range from 0.3 to 0.4 patients per doctor per day. In most provinces, doctors' charges range from ¥20 to ¥30.

Conclusions: Quality doctors are more likely to provide Web-based consultation services, get more patients, and charge higher service fees in an online consultation platform. Policies and promotions could attract more doctors to provide Web-based consultation. The online submarket for the departments of dermatology, psychiatry, and gynecology-obstetrics-pediatrics developed better than other departments such as internal medicine and traditional Chinese medicine. The result could be a reference for policy making to improve the medical system both online and offline. As all the data used for analysis were from a single website, the data might be biased and might not be a representative national sample of China.

(J Med Internet Res 2019;21(6):e12126) doi:10.2196/12126

KEYWORDS

e-consultation; medical service; fee; China

XSL•FO
RenderX

Introduction

Background

Web-based consultation services were first provided to the public by University Hospital of Zurich in 1999 [1]. They have been adopted in many countries to provide better service for patients. An Web-based consultation can serve as an effective complement to traditional health care [2] and conventional physician-patient relationships [3]. Offering Web-based medical services also could make economic sense [4]. Some researchers suggest that 25% to 70% of all patients seeking care do not need a face-to-face appointment with a doctor [5], and there is a great demand for online consultation services.

However, the doctor has been found to be less likely to accept Web-based consultation services than the patient [6]. Consultations, as essentially private information goods, are produced only for a specific consumer [4], and Web-based consultation platforms require a number of doctors to keep patients to attain adequate services. Many of the Web-based consultation platforms are not yet mature and suffer from online doctor scarcity [7]. Furthermore, the Web-based service market was more concentrated on highly skilled doctors' service than the offline service market [8], which makes things worse. Attention needs to be focused on attracting doctors to promote development of Web-based consultation services. Guo et al found that doctors could gain economic and social returns from providing Web-based medical services [9]. Akcura and Ozdemir took an economic perspective to explain doctors adopting Web-based consultation services [4,7,10]. Ozdemir explored doctors' optimal channel strategies via online and offline consultation to get maximum economic profit [4]. Akcura et al examined the conditions under which a quality traditional expert will augment its offline channel by offering an online version of its services [10]. Akcura and Ozdemir investigate when and how expert service providers should offer their services online and what price strategy they should adopt [11]. Akcura and Ozdemir's works are based on game theoretic models and a series of assumptions. However, there is a shortage of description and understanding about doctors who provide Web-based consultation services [7]. In conjunction with the expectation of high-quality internet information, more and more patients are willing to pay for high-quality information provided by specialists online [12]. China is becoming the fastest growing market for Web-based doctor consultation [13]. The market characteristics are also needed to be understood for future development of Web-based consultation services.

Objectives

We tried to fill this gap by investigating doctors' consultation services on Haodf —a Web-based consultation platform in China. The doctors, as the sellers in the Web-based consultation service market, are the key participators we need to focus on. To know more about the doctors, we seek to answer these questions in this paper: (1) which doctors provide Web-based consultation services, (2) how many patients do they serve, and (3) how much do they charge for the Web-based service? We discuss the development of Web-based consultation across departments and provinces. We considered doctors' technical levels, hospitals, departments, and locations to answer the questions. We also compared the condition of doctors' online service with the average level of offline service to understand the online data more exactly.

Methods

Data

As mentioned in the previous section, we obtained the data of doctors who provided Web-based consultation service from Haodf —a Web-based medical service platform. The platform is the *Yelp* for doctor information in China, which lists information of doctors from all over the country and was first built in 2006. Haodf promoted the online review system and the online doctor-patient communication system. Currently, it covers more than half a million doctors for patients to review, according to their own report. Nearly 200,000 doctors have registered and connected with patients through it. Doctors provide Web-based consultation on the website through phones, messages, and graphs in their spare time at a price set by themselves. The rich information about the doctor makes Haodf a perfect platform for doctors to improve reputation. The adoption rate of Haodf as a Web-based consultation platform by doctors is higher than other Web-based consultation platforms [14].

The data of more than 170,000 doctors registered on the platform were collected on March 23, 2017, and September 23, 2017 (we just included doctors in Mainland China, excluding doctors from Hong Kong, Macao, and Taiwan). We removed the data of doctors (1) who did not provide Web-based consultation service when we collected the data; (2) who had not been marked as chief physician, associate chief physician, attending physician, or resident physician; and (3) whose records had missing or abnormal values. We finally got 88,308 doctors who provided Web-based consultation services on the website. The data included a doctor's title, hospital, department, location, services provided, fee for the consultation service, and number of patients served in the past half year. These variables are described in Table 1.

XSL•FO
RenderX

Table 1. Description of the variables.

Variables	Description
Title	Doctor's technical title, values could be chief physician, associate chief physician, attending physician, and resident physician (from senior to junior levels). The title could indicate the doctor's work experience and technical level [15].
Hospital	The hospitals in China are divided into 3 levels by the government according to hospital functions, facilities, and technical strength. Hospitals where the doctor works, values could be level 1 hospital, level 2 hospital, or level 3 hospital (level 3 is the best).
Department	Department, values could be cancer, internal medicine, ophthalmology and otorhinolaryngology, psychiatric, gynecology-obstetrics-pediatrics, dermatology, surgery, traditional Chinese medicine, and others.
Location	The 31 provinces and municipalities in Mainland China, excluding Hong Kong, Macao, and Taiwan.
Online ratio	Number of online doctors/numbers of offline doctors.
Workload	Number of patients a doctor serves per day on Haodf. (We took a doctor's total patient number on March 23, 2017, and September 23, 2017, then, we subtract the previous number from the last one to get the total number of patients in the half year. Finally, we get doctor's total patient number divided by 183.)
Fee	Consulting fee.

Method

Swanson pointed out in 1971 that "*Thinking without comparison is unthinkable. And, in the absence of comparison, so is all scientific thought and scientific research*." To understand the status of Web-based consultation services, we described the doctors' providing Web-based service, patients number, and service price and showed the difference of that across doctors with different technical levels, hospitals, departments, and locations. We also compared consultation services of online doctors and offline doctors. We got the corresponding statistics of doctors providing offline service from the *China Health and Family Planning Statistics Yearbook 2017* [16]. The data include the number of doctors in hospitals; departments and provinces at the end of 2016; as well as the number of patients in hospitals, departments, and provinces in 2016.

To answer "who provides Web-based consultation service," we listed the number of online and offline doctors, and then calculated the online ratio of doctors with different titles, hospital levels, departments, and geographical locations. To answer "how many patients do the Web-based doctors serve," we described the workload distribution on Haodf among doctors with different titles, hospitals, departments, and locations. Then, we made use of the Bonferroni multiple comparison procedures (BMCP) [17] to compare the patient's number of doctors across the subgroups. To exclude the impact of offline doctor-patient distribution among subgroups, we further considered the offline patient number per doctor. With the z test, we compared the online patients' number with adjusted online patients' number per doctor (offline doctor number × average of online patients' number per doctor/average of online patients' number per doctor). As there was no exact offline patient data from March 2017 to September 2017, we adopted the distribution data of patients in 2016.

To answer "what's the price for the services," we described the price distribution among doctors with various titles, different hospitals, different departments, and different geographical positions. Then, we made use of BMCP to compare the price of the service across the subgroups.

Results

Who Provides Web-Based Consultation Service?

We described the characteristic of doctors who provide Web-based consultation and then compare online and offline doctors based on technical level, hospital, department, and location. The results are listed in Table 2 and Figure 1. There are 1,680,062 doctors in China according to the *China Health and Family Planning Statistical Yearbook 2017* [16]. Of these, 5.3% provided Web-based consultation on Haodf as of September 23, 2017.

Referring to doctors' experience and skill level, 15.1% of chief physicians provide Web-based consultation, whereas the online ratio decreases as doctors' experience decreases. The percentages of associate chief physicians, attending physicians, and resident physicians providing Web-based consultations are 8.4%, 5.3% and 2.2%, respectively. Chief physicians and associate chief physicians, who can be called experts, account for nearly 50% of all online doctors, whereas they only account for 25.5% of offline doctors. In the platform, attending physicians get the highest proportion (30,802/88,308, 34.9%) and resident physicians get the lowest proportion (13,440/88,308, 15.2%). In the offline environment, resident physicians get the highest proportion (601,462/1,680,062, 35.8%) and chief physicians get the lowest proportion (119,284/1,680,062, 7.1%).

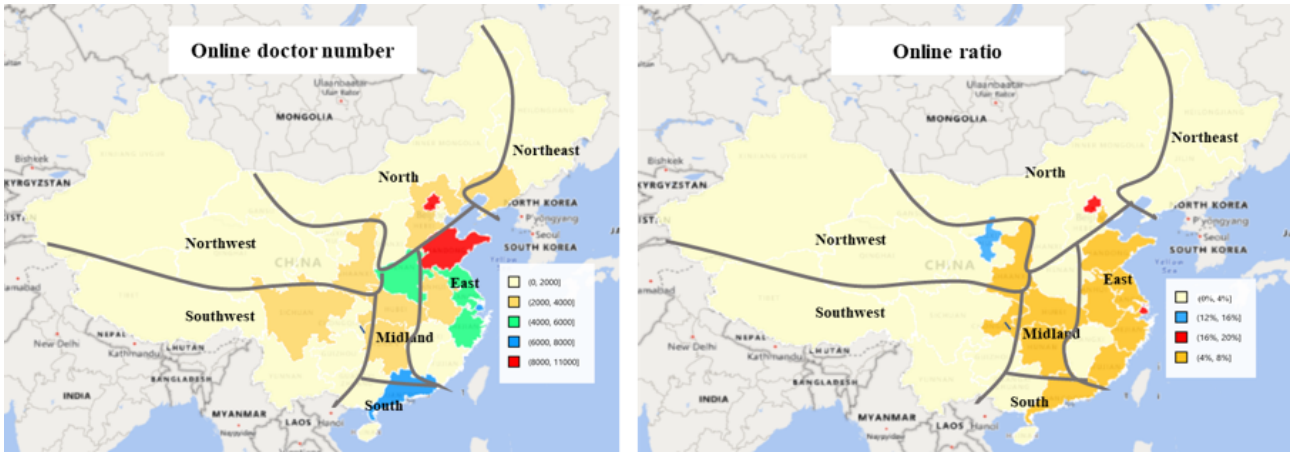
In addition, 9.4% of level 3 hospital doctors provide Web-based consultations, which account for 84.8% of the total number of doctors who provide Web-based consultation. 1.8% of level 2 hospital doctors provide Web-based consultations, which account for 13.0% of the total number of doctors who provide Web-based consultation. However, doctors from level 1 hospitals and no-level hospitals account for just 2.2% of all online doctors.

XSL•FO
RenderX

Table 2. Doctor distribution in the online platform and offline hospitals.

Doctor numbers	Online, n (%)	Offline, n (%)	Online ratio (%)
Total			
All doctors who provide consultation services	88,308 (100.0)	1,680,062 (100.0)	5.3
Title			
Chief physician	18,030 (20.4)	119,284 (7.1)	15.1
Associate chief physician	26,036 (29.5)	309,131 (18.4)	8.4
Attending physician	30,802 (34.9)	581,301 (34.6)	5.3
Resident physician	13,440 (15.2)	601,462 (35.8)	2.2
Unknown	0(0)	68,883 (4.1)	0
Hospital			
Level 3	74,899 (84.8)	795,043 (47.3)	9.4
Level 2	11,454 (13.0)	643,167 (38.3)	1.8
Level 1	718 (0.8)	115,114 (6.9)	0.6
No level	1237 (1.4)	126,738 (7.5)	1.0
Department			
Cancer	2983 (3.4)	18,481 (1.1)	16.1
Internal medicine	16,231 (18.4)	372,974 (22.2)	4.4
Ophthalmology-otorhinolaryngology	8389 (9.5)	132,725 (7.9)	6.3
Psychiatry	1600 (1.8)	16,801 (1.0)	9.5
Gynecology-obstetrics-pediatrics	12,955 (14.7)	225,128 (13.4)	5.8
Dermatology	3334 (3.8)	18,481 (1.1)	18.0
Surgery	24,030 (27.2)	223,448 (13.3)	10.8
Traditional Chinese medicine	8393 (9.5)	223,448 (13.3)	3.8
Others	10,393 (11.8)	448,577 (26.7)	2.3

Figure 1. The doctors' distribution in Mainland China.



Surgery (24,030) and internal medicine (16,231) had the largest number of online doctors. The online ratio of doctors in the dermatology (3,334/18,481, 18.0%), cancer (2,983/18,481, 16.1%), surgery (24,030/223,448, 10.8%), and psychiatry (1,600/16,801, 9.5%) departments providing Web-based consultation are higher than those in other departments. In these 4 departments, doctors in each department account for just 1% of all nationwide doctors, except for surgery (223,448/1,680,062, 13.3%). In the ophthalmology-otorhinolaryngology (8,389/132,725, 6.3%), gynecology-obstetrics-pediatrics (12,955/225,128, 5.8%), internal medicine (16,231/372,974, 4.4%), and traditional Chinese medicine (8,393/223,448, 3.8%) departments, doctors are less likely to provide Web-based services.

Figure 1 shows the number of online doctors and online ratio of doctors in Chinese provinces and cities. Online doctors

XSL•FO
RenderX

mainly come from provinces and cities located in eastern China; in particular, the coastal provinces have more than 4000 online doctors. In most northern and western provinces of China, there are no more than 2000 online doctors. Most provinces and cities in east and midland China have more than 4% of online doctors, which is higher than other regions. Specifically, more than 16% of doctors are online in Beijing and Shanghai. Ningxia Autonomous Region, located in northwest China, also has more than 12% of doctors online.

How Many Patients Does the Online Doctor Serve?

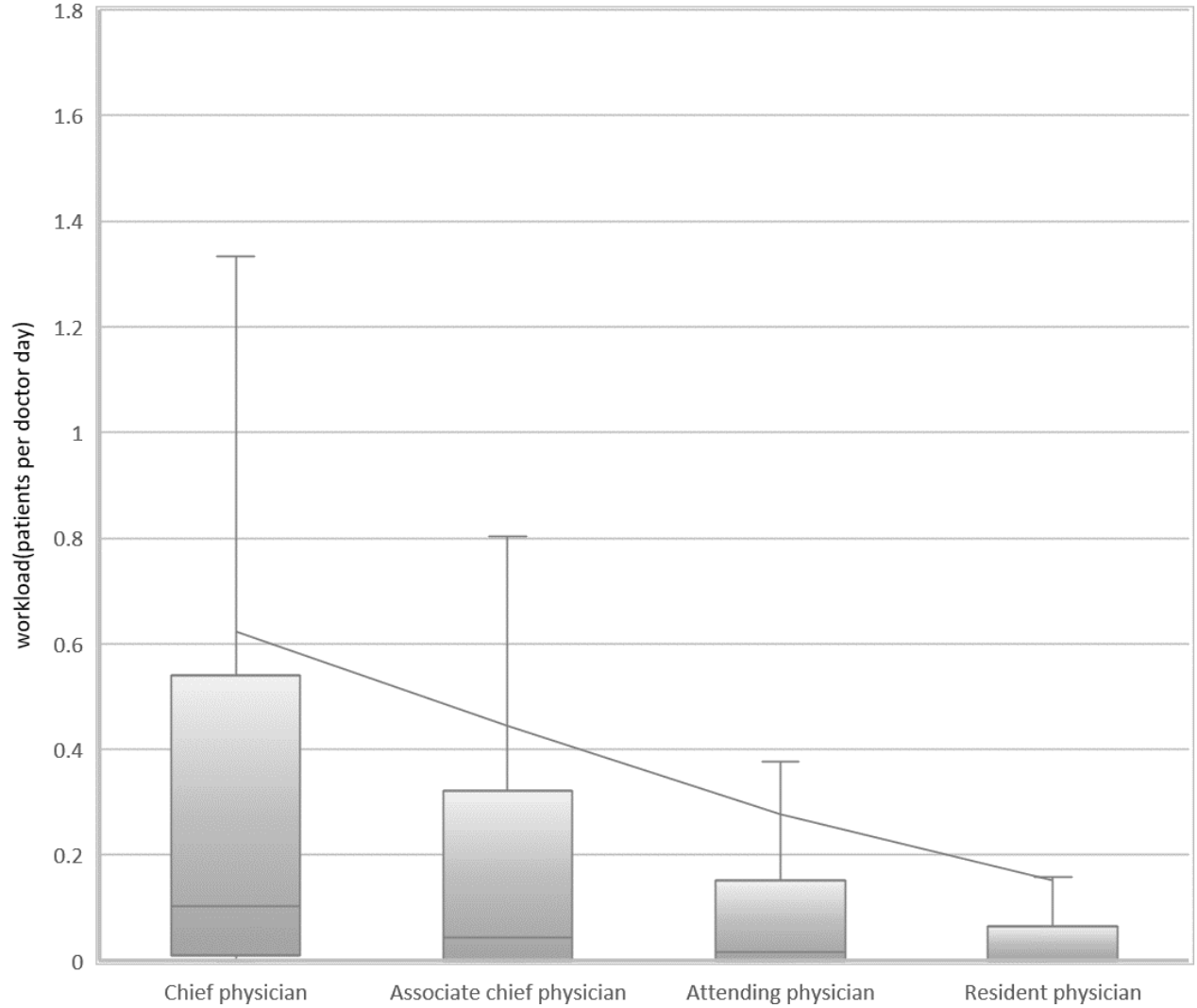
We describe the average patient number a doctor serves in 1 day as the workload for the doctor. Overall, 33.04% of doctors who provided Web-based consultation service got no patients in the past half year. The average online workload of doctors is higher than 75%, which means very few doctors got most of the patients.

We compared the workload of online doctors with that of offline doctors. The average workload of online doctors is 0.38 patients per doctor per day, whereas the average workload of offline doctors is 5.32 patients per doctor per day. The ratio of online doctors' workload to offline doctors is 1:14.

Figure 2 shows the workload distribution of online doctors with different titles, the average workload of chief physicians (0.64), associate chief physicians (0.43), attending physicians (0.28), and resident physicians (0.15) decreased with the decline of doctors' experience and skill level. With BMCP, the online workload for chief physicians is significantly higher than that of associate chief physicians (P<.01), the associate chief physicians are significantly higher than attending physicians (P<.01), the Attending Physicians are significantly higher than Resident Physicians (P<.01).

The workload distribution for online doctors from hospitals is shown in Figure 3. The workloads for doctors from level 3 hospitals are significantly higher than those from level 2 hospitals (P<.01) and level 1 hospitals (P<.01). Thus, high-level hospitals are more attractive to online patients.

Figure 2. Distribution of online and offline patients per doctor across various titles.



XSL•FO
RenderX

Figure 3. Distribution of online and offline patients per doctor across hospitals.

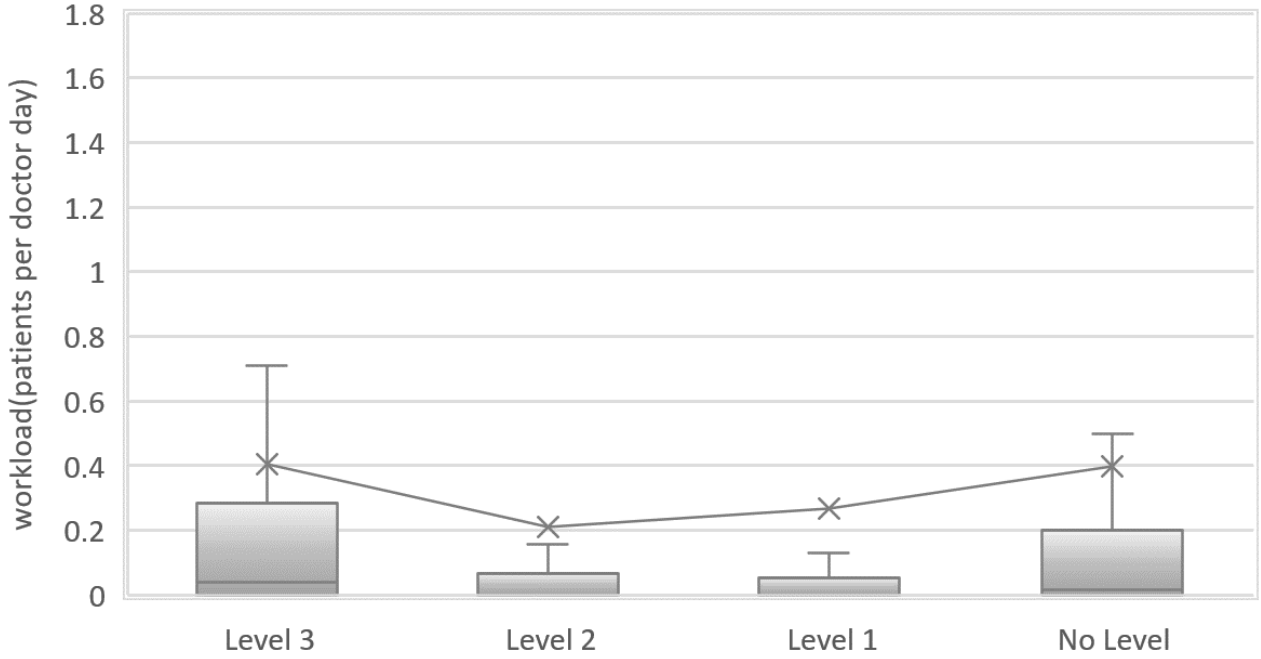


Figure 4 shows the distribution of online workload across departments. With BMCP, dermatology, ophthalmology-otorhinolaryngology, and gynecology-obstetrics-pediatrics are significantly higher than other departments. The workload in psychiatry department is not significantly different from that of most other departments. The online workloads of doctors in the departments of cancer, internal medicine, surgery, and traditional Chinese medicine are lower than the above departments. With z test, the online workload of doctors is compared with the adjusted online workload, ophthalmology-otorhinolaryngology and surgery are significantly higher than the adjusted value, whereas cancer, internal medicine, psychiatry, gynecology-obstetrics-pediatrics, dermatology, and traditional Chinese medicine are significantly lower than the adjusted value.

Figure 5 shows the workload of online and offline doctors in the provinces of China. The offline workloads in Beijing, Shanghai, and most coastal provinces and cities are above 6, whereas the offline workloads in other provinces range from 2 to 6, with no big difference across these provinces. There are 6 provinces with online workloads lower than 0.3 located in the north and west of China. However, the online workloads for most provinces range from 0.3 to 0.4. Beside Beijing and Shanghai, there are also several provinces located in the west of China with online workloads higher than 0.4. In summary, the online workloads are more evenly distributed in cyberspace than offline.

What Is the Price for the Services?

We show the price distribution in Figures 6-8. The average price of a Web-based consultation is ¥32.3. More than 75% of doctors charge no more than ¥32.3—the average fee. The average fee for chief physicians (CFee, ¥56), associate cheif physicians (ACFee, ¥37), attending physicians (AFee, ¥23), and resident physicians (RFee, ¥15) decrease with the decline in experience and skill level. With BMCP, the descending order (CFee > ACFee, ACFee > AFee, AFee > RFee) are significant at .05. The average price for Web-based consultation by doctors from level 3 hospitals is ¥35, twice that of doctors from level 2 and level 1 hospitals.

The price distribution of Web-based medical consultation service varies greatly among different departments. According to the result of BMCP, doctors from psychiatry, gynecology-obstetrics-pediatrics, and dermatology charge more than doctors in other departments. Doctors from traditional Chinese medicine charge an average fee of ¥23, which is lower than any other department.

As shown in Figure 9, doctors from Beijing, Tianjin, and Shanghai charge the highest prices for Web-based consultation services, more than ¥40. In most provinces of eastern and central China, consultation fees range from ¥20 to ¥30. Doctors from western and northern China charge an average fee of no more than ¥20.

XSL•FO
RenderX

Figure 4. Distribution of online and offline patients per doctor across departments.

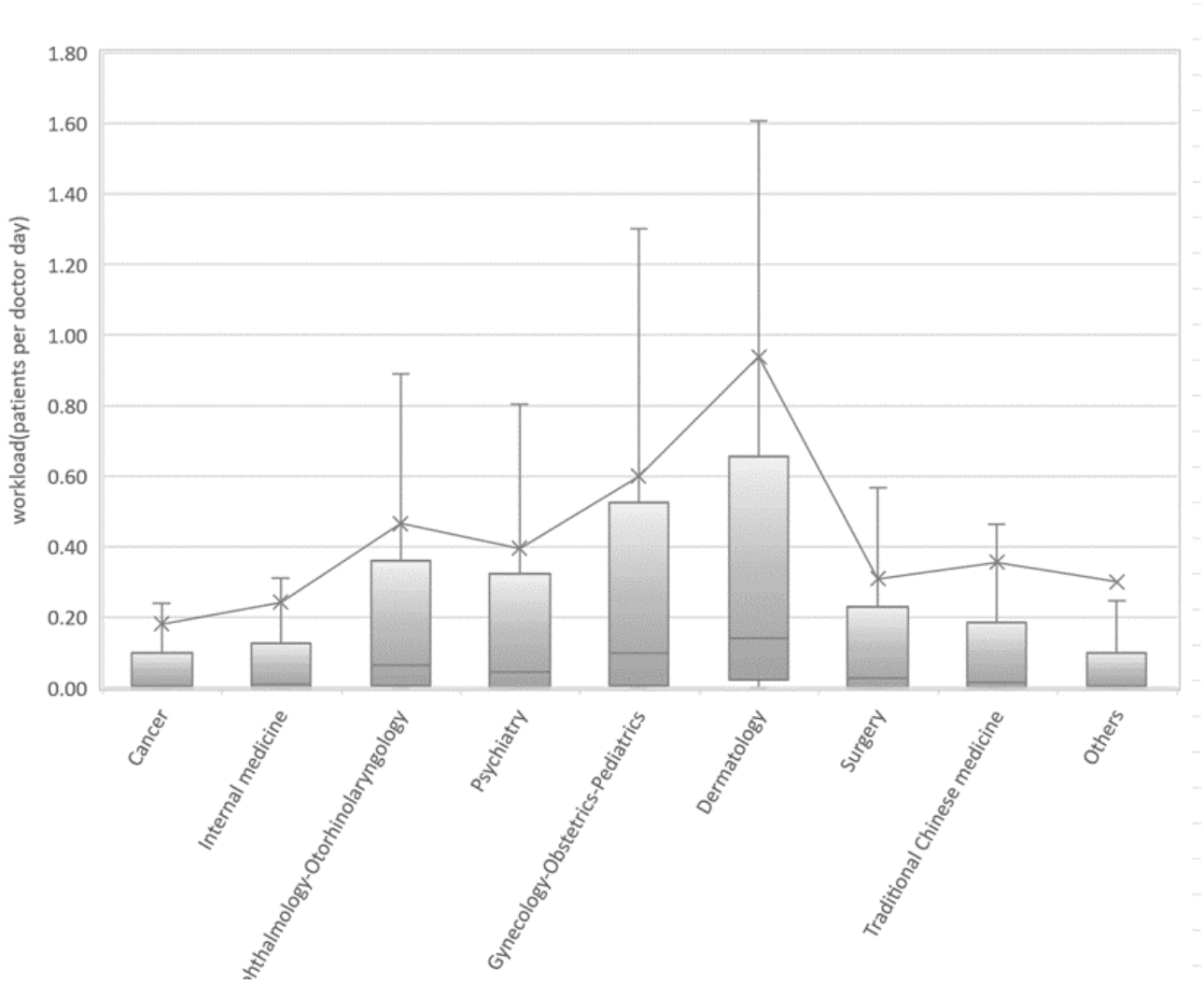
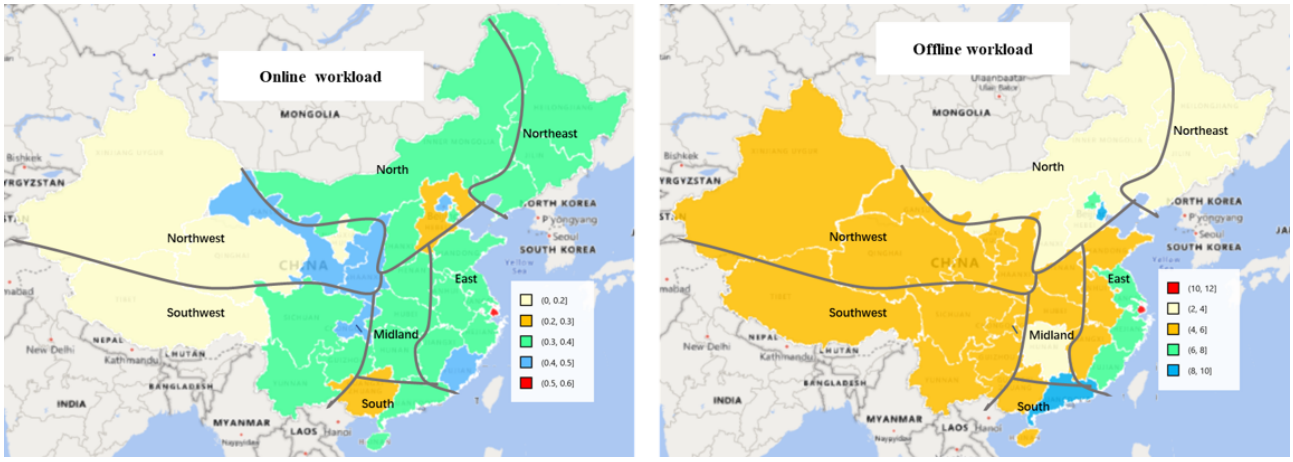


Figure 5. Workload for online and offline doctors in Mainland China.



XSL•FO
RenderX

Figure 6. Distribution of price of Web-based consultations across various doctors' titles.

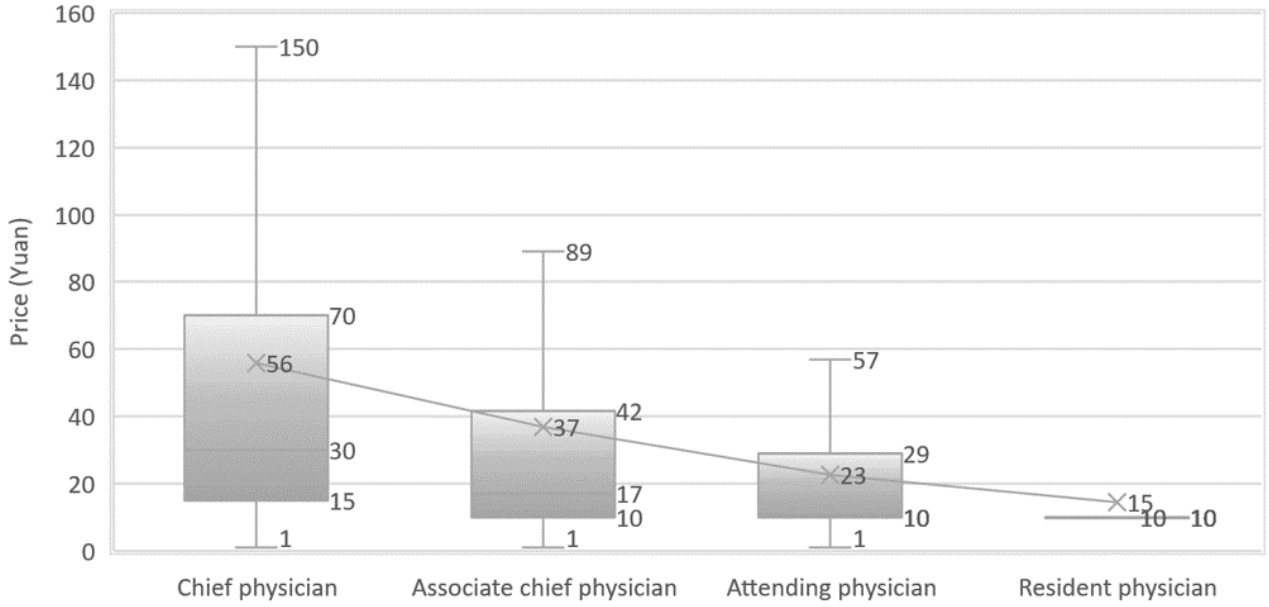
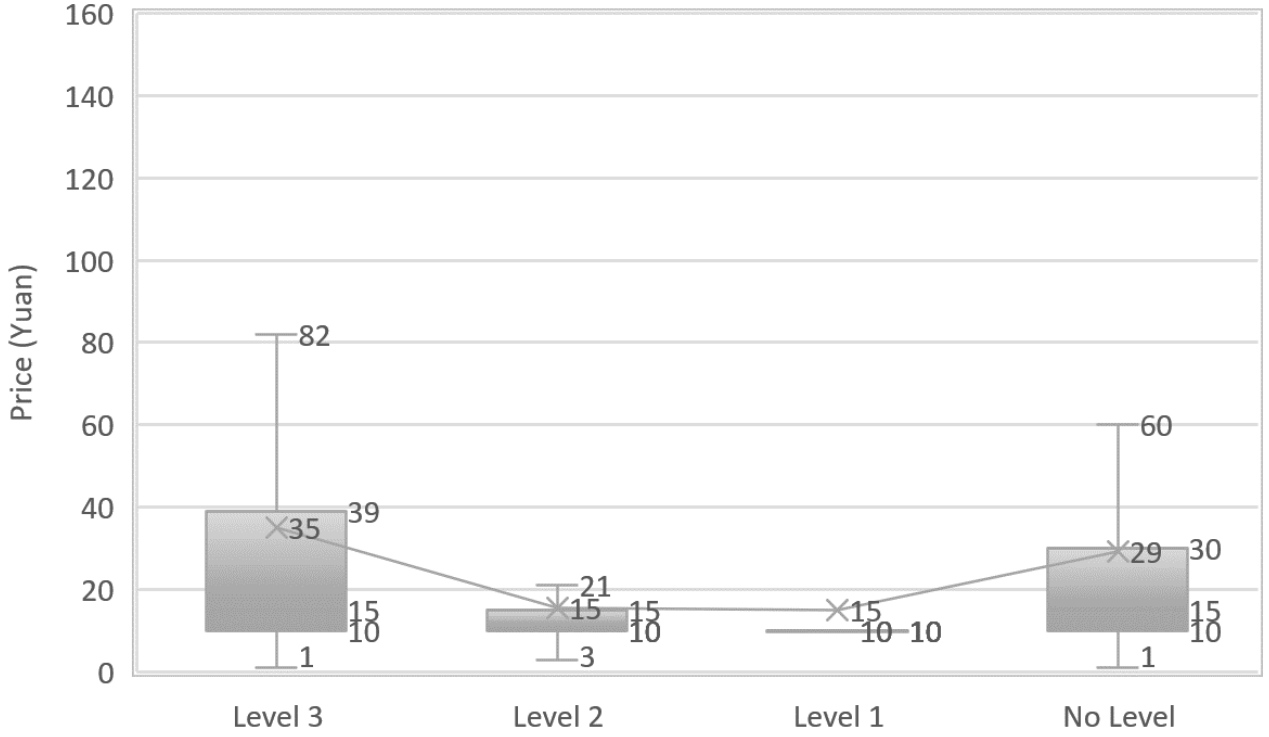
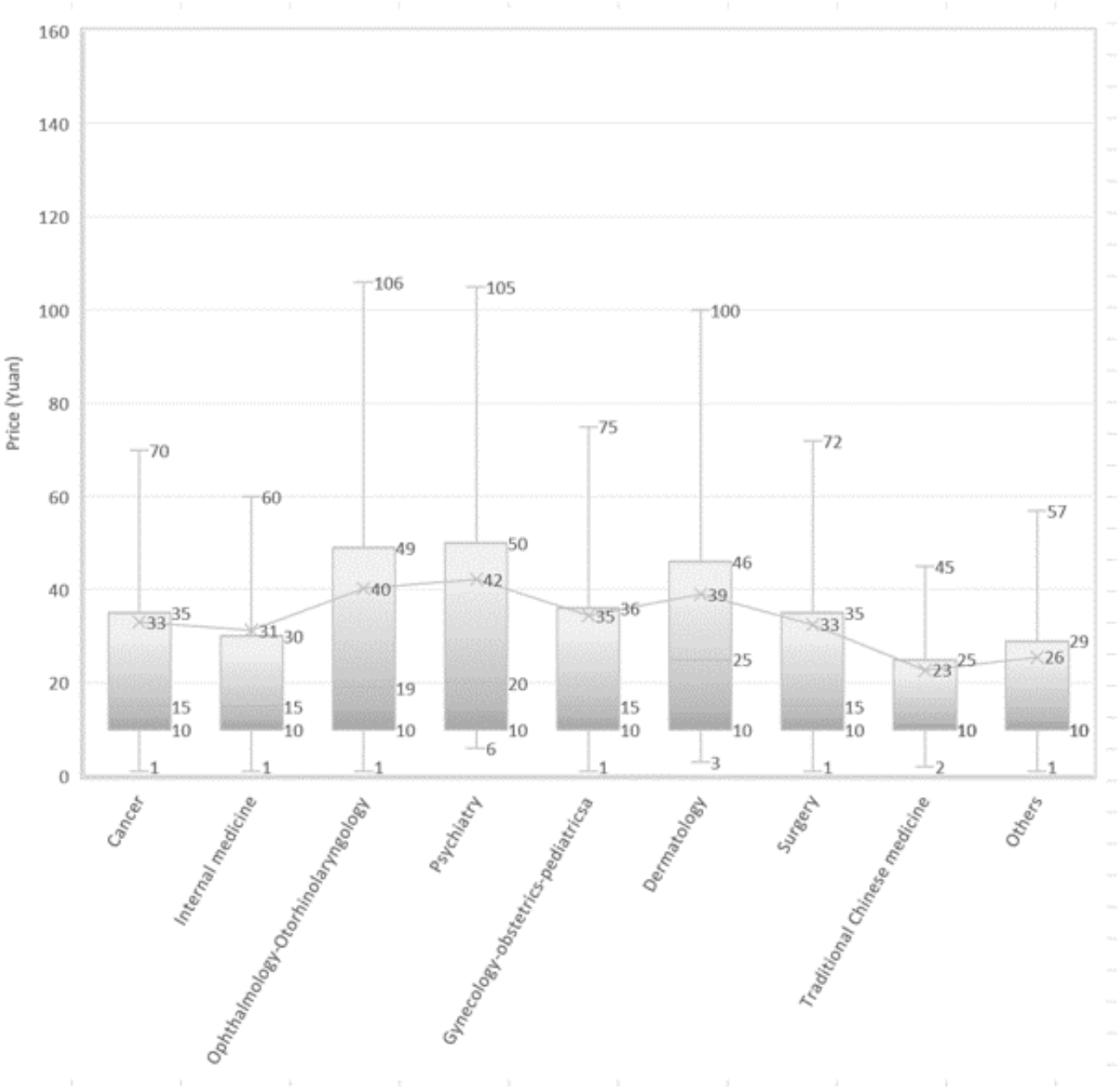


Figure 7. Distribution of price of Web-based consultations across hospitals.



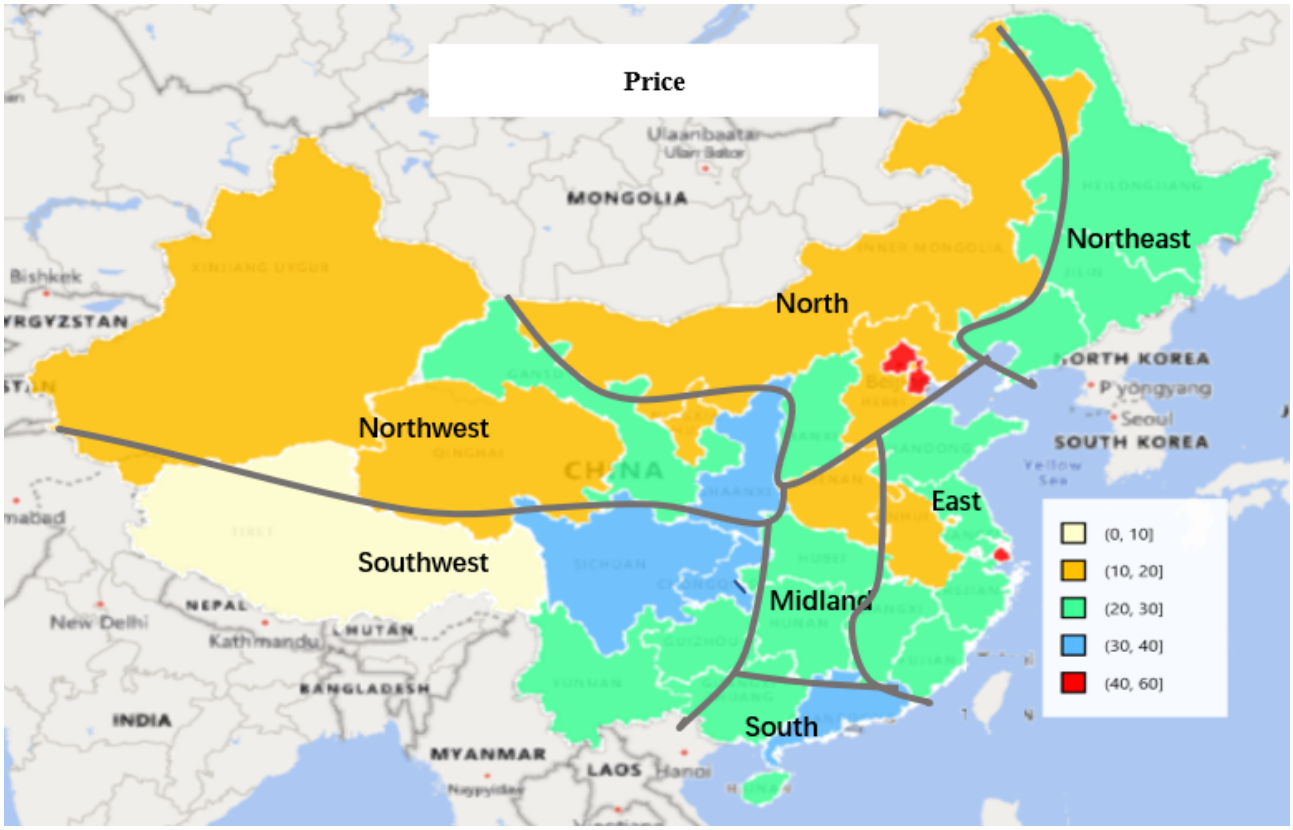
XSL•FO
RenderX

Figure 8. Distribution of price of Web-based consultations across departments.



XSL•FO
RenderX

Figure 9. Distribution of price of Web-based consultations across provinces in Mainland China.



Discussion

Principal Findings

We described the status of doctors' Web-based consultation service on Haodf in September 2017. The platform has become a market for doctors and patients to exchange Web-based consultation service. In China, 6.1% of doctors visit the Web-based consultation platform frequently [14]. There are 88,308 doctors providing Web-based consultation on Haodf in their spare time, accounting for 5.3% of all doctors in China. The adoption rate of Haodf by doctors is higher than of other Web-based consultation platforms [14]. The platform covered doctors from nearly all medical departments all over the country. Therefore, the dataset will assist us to understand the current status of doctors' Web-based consultation in China.

Doctors with senior titles or from top-level hospitals, which are the mark of quality service in both the online and offline environment, are more likely to provide Web-based consultation services on Haodf. Doctors with junior titles or from low-level hospitals are less competitive compared with quality doctors in the Web-based consultation platform, which decreases their positivity toward participation. Furthermore, the quality doctors have higher demand on the internet [8], and the high demand attracts more quality service providers to enter the Web-based consultation market.

Quality doctors are more appealing to patients in the online environment than offline. At the same time, the fees for Web-based consultation service are highest for quality doctors with senior titles or from top-level hospitals. This shows the existence of the Matthew Effect: quality doctors who get more patients and wealth offline than other doctors could get even more resources from an online environment [8]. With patients' increased demand for quality doctors, the platform could help quality doctors to make good use of their spare time to serve more patients. Web-based platforms reduce the cost of searching and accessing quality health care. A large number of quality doctors from all over the country serve on Web-based consultation platforms, providing patients with more options than offline environments. The platform also collects a large amount of information about doctors to provide references for patients, including reputation and experience. With this information, patients can easily find quality doctors from all over the country. Without time and space limitations, patients can communicate with any doctor on the platform through the internet.

Price of the Web-Based Consultation Service

The average price for Web-based consultation service is ¥32, which is higher than the price of offline consultation. The doctor's service is priced mostly by the government and the hospital, especially in the public hospitals, which account for the majority of hospitals in China. For example, the offline consultation price in Anhui is from ¥1 to ¥20 [18] and in Hubei is no more than ¥25 [19,20]. However, the price of Web-based consultation is lower than that of offline consultation in some developed countries. A traditional office visit typically costs US $150 in the United States [21,22] and £100 to £250 in the United Kingdom [23]. Teladoc (from the Unites States) charges US $49 for each consultation, whereas Mdlive (from the United States) charges US $45 per consultation and PushDoctor (from the United Kingdom) charges £20 for 1 consultation. The cost of time could be a reason for the difference. In China, an offline

XSL•FO
RenderX

consultation usually takes about 2 min [24]. Offline consultation usually takes 21 min in the United States and 15 min in the United Kingdom [25]. Web-based consultation could be an unlimited asynchronous message communication or 10 to 20 min through cell phones or videos, which will cost more of the doctor's time than serving an offline patient. In China, the price for offline service is set by the government based on the doctor's title, hospital's level, and location. However, the price of Web-based consultation on the platform ranges from ¥1 to ¥1000. The price of Web-based consultation is set by doctors themselves, which reflects the value of doctors' self-evaluation.

Web-Based Consultation Service in Each Department

There are wide variations in the status of Web-based consultation among doctors in different departments. Table 3 compares the providing of Web-based consultation in each department. In dermatology, ophthalmology-otorhinolaryngology, and gynecology-obstetrics-pediatrics, the online doctor ratio, the online doctor workload, and the price are higher than their average levels, respectively. Doctors in these departments like to provide Web-based services and sell their service at a higher price. This shows good development of a submarket in these departments. In dermatology, the visual examination of the skin is often the key part of the consultant's physical examination [26]. The visual nature makes dermatology an obvious candidate for telemedicine techniques, and the feasibility and reliability of teledermatology is already well established [27]. The well-developed online market for gynecology-obstetrics-pediatrics should result from the high demand for quality service in this department. There is a shortage of doctors for this department in China [28], which leads to overcrowding in hospitals. The less crowded online consultation platform could be an alternative choice. Although children are always the focus of a family and the extended family in Chinese culture, an Web-based consulting service that is not limited by time or space is ideal if an urgent need arises. Web-based consultation in psychiatry has occurred since the mid-1990s [29]. There are no known absolute exclusion criteria or contraindications for any specific psychiatric diagnoses, treatments, or populations [30]. Psychiatry is an ideal fit for Web-based consultation [30]. In traditional Chinese medicine, compared with the high offline patient number per doctor, the low online doctor ratio, online workload, and price show this is a poorly developed online submarket. Traditional Chinese medicine practitioners prefer the 4 ways of diagnosis— "look, listen, question and feel the pulse" [31], which are difficult to do in the online environment. In cancer and surgery, the high online ratio and the low online workload show the excess supply of doctors and low demand in these 2 departments. High remuneration in these 2 departments leads medical students to choose these specialties [32], which decreases the patient numbers for each doctor. Providing Web-based consultations could give doctors an optimal opportunity to access more patients. Furthermore, Web-based consultations are a good opportunity to give or get a second opinion. In internal medicine, online doctor ratio, online patient number per doctor, and the price are lower than the average values. To cope with the aging population and increase in chronic diseases, the government has increased the allocation of medical resources in internal medicine, which temporarily leads to more doctors and fewer patients per doctor. The low online ratio may be because some physical examination, which is the basis for diagnosis, is not possible on the platform.

Table 3. Comparison of Web-based consultation provision in each department.

Department	Online ratio	Price	Online workload	Adjusted online workload
Cancer	+[a]	+	–[b]	–
Internal medicine	–	–	–	–
Ophthalmology-otorhinolaryngology	+	+	+	–
Psychiatry	+	+	N[c]	–
Gynecology-obstetrics-pediatrics	+	+	+	+
Dermatology	+	+	+	–
Surgery	+	+	–	+
Traditional Chinese medicine	–	–	–	–
Others	–	–	–	–

[a]The value is significantly higher than the average value of all online doctors.

[b]The value is significantly lower than the average value of all online doctors.

[c]The value has no significant difference from the average value of all online doctors.

Web-Based Consultation Service in Provinces

Beijing and Shanghai have higher online ratio, workload, and fees for Web-based consultations than other provinces. As the economic centers of China, Beijing and Shanghai have more top hospitals [33] and more famous doctors [34]. The quality medical resources attract patients from all over the country. Without transportation cost, Web-based consultations become an excellent way to access the quality medical service in Beijing and Shanghai. Quality doctors and easy access lead to the high demand for Web-based service of doctors in these places. The quality and large demand result in higher fees for the Web-based service of these doctors. The high demand and high remuneration encourage the doctors' participation, which shows

XSL•FO
RenderX

as the high online ratio. Furthermore, having centralized quality doctors leads to fierce competition among doctors in Beijing and Shanghai. Reputation is a good way to stand out from the crowd, and Web-based service is a good way to build reputation. The Web-based platform can be an optional channel for doctors to communicate with offline patients as well.

The workloads are more evenly distributed across provinces in the online environment. This may be because of the platform's triage and market power. The platform has a large number of online doctors from all over the country and shares their detailed information. The platform helps patients find the right doctor nearby rather than contending for quality doctors who are overbooked. The platform could balance the workload among all its doctors with appropriate professional competence. Compared with the offline market, which is affected by the government's macroplan, the online market is easier to enter and leave. Doctors could become an online doctor by connecting to the internet. The internet covers 55.8% of people in China, more in cities [35], and doctors are more likely to adopt the internet. Nearly all offline doctors are potential online market entrants. Doctors could simply choose to enter or leave the market according to the market size.

It is well to remember that Ningxia Autonomous Region, located in west China where the economy is less developed and medical resources are scarce, has a higher percentage of doctors online than all other provinces except Beijing and Shanghai. This could be caused by the policies and the government's intervention. The local government of Ningxia Autonomous Region published *Yinchuan Internet Hospital Management Working System (trial)* and *Yinchuan Internet Medical Institution Supervision and Management System (trial)* in August 2016. They are the first internet hospital regulations in China. The promotion policies are directly related with the doctor's income and attract doctors to provide online consultation. At the same time, the online workload and service fee are very low compared with other provinces. This shows that policies could guide doctor participation although they may not attract patients.

Conclusions

There are 88,308 doctors providing Web-based consultation in their spare time on Haodf as of September 23, 2017, accounting for 5.3% of all doctors in China. There are 0.38 patients per online doctor per day on average. The average price of Web-based consultation is ¥32.3, which is higher than the price for offline consultation. The online submarkets for dermatology, psychiatry, and gynecology-obstetrics-pediatrics are better developed than for other departments such as internal medicine and traditional Chinese medicine. Government policy could promote doctors' participation in Web-based consultation.

The first contribution of this paper is that it is the first study, or one of the first, to examine the status of the provision of Web-based consultation in China. This study explores the status of doctors in each department and location providing Web-based consultation services. This paper could be a reference for Web-based medical service providers to establish Web-based consultation services, for doctors to make decisions about Web-based consultation, and for government policy making.

Limitations

This research has some limitations. First, we included just one of the Web-based consultation platforms, although it is the most acceptable one. Second, we only have the aggregate data of all offline doctors in China instead of each doctor's detailed information.

Acknowledgments

This research was supported in part by the Research Funds of the National Natural Science Foundation of China (Grant Numbers: 71531013, 71601034) and the China Postdoctoral Science Foundation (Grant Number: 2016M600210).

Conflicts of Interest

None declared.

References

1. Schmidt CA, Schmidt-Weitmann SH, Lachat ML, Brockes CM. Teleconsultation in vascular surgery: a 13 year single centre experience. J Telemed Telecare 2014 Jan;20(1):24-28. [doi: 10.1177/1357633X13517356] [Medline: 24352901]
2. Lu H, Shaw BR, Gustafson DH. Online health consultation: examining uses of an interactive cancer communication tool by low-income women with breast cancer. Int J Med Inform 2011 Jul;80(7):518-528 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.03.011] [Medline: 21530381]
3. Brockes C, Schenkel JS, Buehler RN, Grätz K, Schmidt-Weitmann S. Medical online consultation service regarding maxillofacial surgery. J Craniomaxillofac Surg 2012 Oct;40(7):626-630. [doi: 10.1016/j.jcms.2012.03.018] [Medline: 22555239]
4. Ozdemir ZD. Optimal multi-channel delivery of expertise: an economic analysis. I J Electron Commer 2014 Dec 08;11(3):89-105. [doi: 10.2753/jec1086-4415110303]
5. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [Medline: 15076819]
6. Bashshur RL, Howell JD, Krupinski EA, Harms KM, Bashshur N, Doarn CR. The empirical foundations of telemedicine interventions in primary care. Telemed J E Health 2016 May;22(5):342-375 [FREE Full text] [doi: 10.1089/tmj.2016.0045] [Medline: 27128779]

XSL•FO
RenderX

7. Xie X, Zhou W, Lin L, Fan S, Lin F, Wang L, et al. Internet hospitals in China: cross-sectional survey. J Med Internet Res 2017 Jul 04;19(7):e239 [FREE Full text] [doi: 10.2196/jmir.7854] [Medline: 28676472]
8. Li J, Zhang Y, Ma L, Liu X. The impact of the internet on health consultation market concentration: an econometric analysis of secondary data. J Med Internet Res 2016 Oct 28;18(10):e276 [FREE Full text] [doi: 10.2196/jmir.6423] [Medline: 27793793]
9. Guo S, Guo X, Fang Y, Vogel D. How doctors gain social and economic returns in online health-care communities: a professional capital perspective. J Manage Inf Syst 2017 Aug 17;34(2):487-519. [doi: 10.1080/07421222.2017.1334480]
10. Akçura MT, Ozdemir ZD, Jain S. Expert Competition and the Internet. International Journal of Electronic Commerce 2014 Dec 06;18(1):11-44. [doi: 10.2753/jec1086-4415180101]
11. Akçura MT, Ozdemir ZD. A strategic analysis of multi-channel expert services. J Manag Inf Sys 2017 Apr 20;34(1):206-231. [doi: 10.1080/07421222.2017.1297637]
12. Analysys. 2016. The White Paper on the Development of China Knowledge Payment Industry 2016 URL: https://www.analysys.cn/article/analysis/detail/1000416 [accessed 2019-05-19] [WebCite Cache ID 78UMIDM7E]
13. Garner Insights. Online Doctor Consultation Market: Global Market Synopsis, Growth Factors, Industry Segmentation, Regional Analysis And Competitive Analysis 2017 - 2025 URL: http://garnerinsights.com/Online-Doctor-Consultation-Market-Global-Market-Synopsis-Growth-Factors-Industry-Segmentation-Regional-Analysis-And-Competitive-Analysis-2017--2025 [accessed 2019-05-19]
14. IResearch. Report of China Internet medical doctors' needs in 2016 URL: http://report.iresearch.cn/report/201609/2650.shtml [accessed 2019-05-19] [WebCite Cache ID 78Tz13BFO]
15. Hao H. The development of online doctor reviews in China: an analysis of the largest online doctor review website in China. J Med Internet Res 2015 Jun 01;17(6):e134 [FREE Full text] [doi: 10.2196/jmir.4365] [Medline: 26032933]
16. China Health and Family Planning Statistics. 2017. National Health and Family Planning Commission of the People's Republic of China URL: http://www.tjcn.org/tjnj/www/36307.html [accessed 2019-05-19] [WebCite Cache ID 78UKJAX6H]
17. Savin NE. The Bonferroni and the Scheffe multiple comparison procedures. Rev Econ Stud 1980 Jan;47(1):255. [doi: 10.2307/2297112]
18. Anhui Price Bureau. 2018. Medical Service Price in Anhui URL: http://www.ahpi.gov.cn/DocHtml/1/16/05/xxgk_201607215341764.html [accessed 2019-05-19] [WebCite Cache ID 78UKReJgF]
19. Hubei Price Bureau. 2018. Price adjustment from medical reform in 2017 (other public hospitals in Wuhan) URL: http://www.hbpic.gov.cn/jgggfw/sfgl/ylsf/201801/t20180123_32630.html [accessed 2019-05-19] [WebCite Cache ID 78UKH8EXu]
20. Hubei Price Bureau. 2018 Aug 01. Price adjustment from medical reform in 2017 (Wuhan tongji hospital, Wuhan Union Hospital) URL: http://www.hbpic.gov.cn/jgggfw/sfgl/ylsf/201801/t20180123_32620.html [accessed 2019-05-19] [WebCite Cache ID 78UKa3dvL]
21. US News. 2012. Is Your Doctor Out of the Office? Try an E-Visit URL: https://health.usnews.com/health-news/articles/2012/09/04/is-your-doctor-out-of-the-office-try-an-e-visit [accessed 2019-05-19] [WebCite Cache ID 78TzLQBR8]
22. The New York Times. 2015. Modern Doctors' House Calls: Skype Chat and Fast Diagnosis URL: https://www.nytimes.com/2015/07/12/health/modern-doctors-house-calls-skype-chat-and-fast-diagnosis.html [accessed 2019-05-19] [WebCite Cache ID 78U7rcGhE]
23. Netdoctor. How much will a private consultant charge to see me? URL: https://www.netdoctor.co.uk/health-services/private-treatment/a4556/how-much-will-a-private-consultant-charge-to-see-me/ [accessed 2019-05-19] [WebCite Cache ID 78U84CI09]
24. Jin G, Zhao Y, Chen C, Wang W, Du J, Lu X. The length and content of general practice consultation in two urban districts of Beijing: a preliminary observation study. PLoS One 2015;10(8):e0135121 [FREE Full text] [doi: 10.1371/journal.pone.0135121] [Medline: 26258911]
25. Centers for Disease Control and Prevention. 2012. National Ambulatory Medical Care Survey: 2012 State and National Summary Tables URL: https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2012_namcs_web_tables.pdf [accessed 2019-05-19] [WebCite Cache ID 78ULmuS3m]
26. Bowns IR, Collins K, Walters SJ, McDonagh AJ. Telemedicine in dermatology: a randomised controlled trial. Health Technol Assess 2006 Nov;10(43):iii-iv, ix [FREE Full text] [Medline: 17049140]
27. Soyer HP, Hofmann-Wellenhof R, Massone C, Gabler G, Dong H, Ozdemir F, et al. telederm.org: freely available online consultations in dermatology. PLoS Med 2005 Apr;2(4):e87 [FREE Full text] [doi: 10.1371/journal.pmed.0020087] [Medline: 15839749]
28. Chunju L, Li C. Causes and countermeasures of shortage of pediatricians in China. The Journal of Medical Theory and Practice 2017:2181-2183.
29. 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]
30. Shore JH. Telepsychiatry: videoconferencing in the delivery of psychiatric care. Am J Psychiatry 2013 Mar;170(3):256-262. [doi: 10.1176/appi.ajp.2012.12081064] [Medline: 23450286]
31. Li X. Analysis of the clinical effect of Chinese medicine herb in the treatment of diabetes. Chinese Community Doctors 2014:94-95.

XSL•FO
RenderX

32. Ming Z. International experience and enlightenment in the design of doctor pay system. Finance and Economics (academic edition) 2016:358.
33. Fundamed. 2016. Top 100 hospitals ranking in China URL: http://www.fudanmed.com/institute/news2016-2.aspx [accessed 2019-05-19] [WebCite Cache ID 78UJQaa1H]
34. Huayizonghengbang. 360doc. 2017. The most influential doctors rankings in China URL: http://www.360doc.com/content/17/0401/21/31429017_642177204.shtml [accessed 2019-05-19] [WebCite Cache ID 78UJUuLu9]
35. The China Internet Network Information Center. 2018. 41th Statistical Report on Internet Development in China URL: http://www.cac.gov.cn/2018zt/cnnic41/index.htm [accessed 2019-05-19] [WebCite Cache ID 78UJgsTJ8]

Abbreviations

BMCP: Bonferroni multiple comparison procedure

Edited by G Eysenbach; submitted 07.09.18; peer-reviewed by X Xie, X Shen, A Sule; comments to author 14.10.18; revised version received 23.03.19; accepted 26.04.19; published 03.06.19.

Please cite as:
Li Y, Yan X, Song X
Provision of Paid Web-Based Medical Consultation in China: Cross-Sectional Analysis of Data From a Medical Consultation Website
J Med Internet Res 2019;21(6):e12126
URL: https://www.jmir.org/2019/6/e12126/
doi:10.2196/12126
PMID:31162129

©Yumei Li, Xiangbin Yan, Xiaolong Song. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Telemedicine for Patients With Inflammatory Bowel Disease (TELE-IBD) Clinical Trial: Qualitative Assessment of Participants' Perceptions

Charlene Connolly Quinn[1*], RN, PhD; Sarah Chard[2*], PhD; Erin G Roth[2], PhD; J Kevin Eckert[2], PhD; Katharine M Russman[3*], BA; Raymond K Cross[3*], MD, MS

[1]University of Maryland School of Medicine, Department of Epidemiology and Public Health, Baltimore, MD, United States

[2]University of Maryland Baltimore County, Department of Sociology, Anthropology, and Health Administration and Policy, Baltimore, MD, United States

[3]University of Maryland School of Medicine, Department of Medicine, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Charlene Connolly Quinn, RN, PhD
University of Maryland School of Medicine
Department of Epidemiology and Public Health
660 W Redwood St, HH 215
Baltimore, MD, 21201
United States
Phone: 1 410 706 2406
Fax: 1 410 706 4433
Email: cquinn@som.umaryland.edu

Abstract

Background: Inflammatory bowel disease (IBD), comprising Crohn disease and ulcerative colitis, affects 1 to 3 million people in the United States. Telemedicine has shown promise in IBD. The objective of this study, telemedicine for patients with IBD (TELE-IBD), was to compare disease activity and quality of life (QoL) in a 1-year randomized clinical trial of IBD patients receiving telemedicine versus standard care. Treatment groups experienced improvements in disease activity and QoL, but there were no significant differences between groups. Study adherence to the text-based intervention was less than 80%, the targeted goal.

Objective: To understand adherence to remote monitoring, the goal of this qualitative assessment was to obtain TELE-IBD trial participants' perceptions, including their recommendations for future monitoring.

Methods: In this study, patients attending 3 tertiary referral centers with worsening IBD symptoms in the previous 2 years were eligible for randomization to remote monitoring via SMS text messages (short message service, SMS) every other week, weekly, or standard care. Participants (n=348) were evenly enrolled in the treatment groups, and 259 (74.4%) completed the study. For this study, a purposive sample of adherent (N=15) and nonadherent (N=14) patients was drawn from the TELE-IBD trial population. Adherence was defined as the completion of 80% (278/348) or more of the weekly or every other week self-assessments. Semistructured interviews conducted by phone surveyed (1) the strengths and benefits of TELE-IBD, (2) challenges associated with using TELE-IBD, and (3) how to improve the TELE-IBD intervention. Interviews were recorded, professionally transcribed, and coded based on *a priori* concepts and emergent themes with the aid of ATLAS.ti, version.7 qualitative data analysis software.

Results: Participants' discussions centered on 3 elements of the intervention: (1) self-assessment questions, (2) action plans, and (3) educational messages. Participants also commented on text-based platform, depression and adherence, TELE-IBD system in place of office visit, and their recommendations for future TELE-IBD systems. Adherent and nonadherent participants prefer a flexible system that is personalized, including targeted education messages, and they perceive the intervention as effective in facilitating IBD self-management.

Conclusions: Participants identified clear benefits to the TELE-IBD system, including obtaining a better understanding of the disease process, monitoring their symptoms, and feeling connected to their health care provider. Participants' perceptions obtained in this qualitative study will assist in improving the TELE-IBD system to be more responsive to patients with IBD.

XSL•FO
RenderX

(J Med Internet Res 2019;21(6):e14165) doi:10.2196/14165

KEYWORDS

inflammatory bowel diseases; Crohn disease; ulcerative colitis; qualitative research; telemedicine

Introduction

Inflammatory bowel disease (IBD), comprising Crohn disease (CD) and ulcerative colitis (UC) affects nearly 1 to 3 million people in the United States [1,2]. IBD are chronic diseases characterized by abdominal pain, diarrhea, bloody stools, fatigue, and extraintestinal manifestations [2]. The symptoms have a negative impact on quality of life (QoL) [3] and result in increased health care utilization [4-6]. Effective treatments exist for patients with IBD. However, a significant number of patients have suboptimal outcomes with standard care. Reasons for suboptimal outcomes include nonadherence [5,7], delays initiating treatment, inadequate monitoring, side effects [8,9], poor education [10,11], and lack of access to IBD care [12]. Telemedicine is increasingly being evaluated by health care systems and payers as an alternative service to address deficiencies in health care delivery [13,14].

The objective of this study, telemedicine for patients with IBD (TELE-IBD), was to compare disease activity and QoL in a 1-year randomized clinical trial of IBD patients receiving telemedicine versus standard care. Treatment groups experienced improvements in disease activity and QoL, but there were no significant differences between groups. Study adherence to the text-based intervention was less than 80%, the targeted goal [13]. To understand treatment adherence, the goal of the qualitative assessment reported here was to obtain TELE-IBD trial participants' perceptions of the system, including their recommendations for future TELE-IBD monitoring.

Methods

Study Design, Setting, and Recruitment

Parent Randomized Controlled Trial

Study trial protocol and results of the parent randomized clinical study have been previously published [13,15]. Briefly, the TELE-IBD study was a multicenter, randomized, controlled, clinical trial conducted over 12 months. Participants were adults (mean 38.9, SD 12.3 years; 56.6% women) recruited from 3 academic medical center IBD specialty clinics. Participants in the intervention arms completed self-testing with TELE-IBD on either a weekly or every other week basis using a mobile phone–based platform. The system worked on both iPhone operating system and Android operating systems. All but 3 patients used their own phone. A total of 3 patients received a mobile phone from the study for study trial use only.

Participants randomized to the control group received standard care. TELE-IBD was designed using a mobile phone for participants and a decision support server with website for staff and providers. The website provided an interface for staff and providers for participant profiles and collected data from each self-testing session. The provider could individualize alerts and action plans for each participant, if needed. A standard set of baseline action plans based on the criteria were available to providers. If predetermined criteria were met after testing, simultaneous action plans and email alerts were sent to the participant and study team, respectively. The study team reviewed the information and if necessary, consulted the prescribing provider for management changes. Medication changes were updated in the participant profile and communicated to the participant.

Participants in the 2 intervention groups were prompted to respond to a series of SMS text messages (short message service, SMS) grading their IBD symptoms. To assess bowel symptoms and extraintestinal manifestations, abbreviated disease activity forms were used. CD participants completed a modified Harvey Bradshaw Index [16,17]. Overall well-being, abdominal pain, and diarrhea were core questions. UC and indeterminate colitis (IC) participants completed the Simple Clinical Colitis Activity Index [18,19]. Overall well-being, number of bowel movements during the day, and presence of blood in the stool were core questions.

After answering SMS text message questions about their symptoms, participants received a list of medication(s), dose(s), and direction(s). Participants were asked if they experienced any drug side effects. Participants were asked to describe and grade the severity (mild, moderate, or severe) of side effect; moderate to severe side effects generated an alert.

Alerts and action plans were customized for participants and generated based on responses to the core questions in the activity indices for CD, UC, and IC. Depending on the response, participants were assigned to a disease activity zone: the green zone represented patient remission or mild disease activity, the yellow zone represented moderate disease activity, and the red zone represented severe disease activity. Depending on the severity zone, the provider could select an action for the participant to begin.

TELE-IBD intervention participants received an educational tip twice weekly (W group) or every week (EOW group) and at the discretion of the provider. For example, the provider could send an SMS text message during flu season asking participants to obtain a flu vaccination.

Qualitative Assessment of the Telemedicine for Patients With Inflammatory Bowel Disease Intervention

Participants, Design, and Methods

For the descriptive qualitative assessment reported here, the purposive sample of adherent (N=15) and nonadherent (N=14) patients was drawn from the TELE-IBD trial population. To achieve this sample, 50 patients were contacted from each group of adherent and nonadherent patient participants. No patients refused out of disinterest; 1 patient was out of the country, and the remaining patients could not be contacted after multiple attempts. The final sample was not different from those who participated in this qualitative study. Adherence was defined as

XSL•FO
RenderX

the completion of 80% or more of the weekly or every other week self-assessments; 3 core questions described above for each IBD condition had to be answered for the self-assessment to be considered adherent. Participants were contacted via telephone; interviewers described the qualitative study and determined participant willingness to participate in this follow-up study. Phone interviews were scheduled at the convenience of participants; the study was described again, and verbal informed consent was obtained. The study Interview Guide is included in the supplementary file as a Multimedia Appendix 1.

Data Analysis

Qualitative analysis was used to identify themes of participants' interviews. A codebook of 24 codes was developed based on *a priori* concepts of interest, including symptoms, TELE-IBD core questions, action plans, education messages, and medications and themes that emerged from the SMS text message. Braun and Clarke's 6 phases of thematic analysis was applied to organize and identify core themes and subthemes [20]. This approach included several steps by the research team including familiarization, professional transcription, generating initial codes, searching for themes, reviewing themes, and finally defining and deciding on meaningful themes. An iterative approach was adopted with the initial coding performed by ER and SC following the completion of the first set of 10 interviews (5 adherent, 5 nonadherent). Subsequent interviews were then conducted and transcribed incrementally with data analysis occurring simultaneously, and new codes were added from the dataset to ATLAS.ti, version 7. ER, SC, and KE reviewed the transcripts and codes during this process until consensus was reached by all members of the research team. Pattern saturation was achieved (and data collection ceased) when no new themes emerged, determined by the point in analysis where no new codes would provide additional value to the identified themes.

A qualitative research team member coded each interview in ATLAS.ti, version 7. A second team member independently reviewed the coding for missed codes or alternative interpretations. Discrepancies were discussed between coders. The process involved sorting and sifting through the narratives to confirm coding and resolve discrepancies in interpretation and meaning. Coded narrative was then compared across the adherent and nonadherent participants. On the basis of comparative analysis, coded narratives were used to develop higher order themes related to their meanings. Team members then reviewed individual transcripts to identify how themes were embedded within the complete interviews. This served to reach consensus on the narrative meaning. Finally, codes were analyzed to determine patterns within and across the adherent and nonadherent participants. Team members also reviewed individual transcripts to identify how themes were embedded within the complete interview. This review served as an important qualitative validity check.

Interviews lasted between 18 to 54 min and examined perceptions of (1) the strengths and benefits of the text-based intervention (TELE-IBD system), (2) challenges associated with using the TELE-IBD, and (3) how to improve the system. The study protocol was reviewed and approved by the Human Research Protection Office at the University of Maryland, Baltimore.

Measures

Interview guide questions created by study team members were based on (1) the strengths and benefits of TELE-IBD, (2) challenges associated with using TELE-IBD, and (3) how to improve the TELE-IBD intervention (see supplementary materials).

Results

Participant Characteristics

Adherent (n=15) and nonadherent (n=14) participants completed interviews (see Table 1). Participants represented patients with both CD (n=16) and UC (n=12) and 1 patient with IC. Both intervention groups were represented, although the majority were participants receiving every other week SMS text messages (69%).

XSL•FO
RenderX

Table 1. Demographic characteristics of participants from the telemedicine for patients with inflammatory bowel disease trial participating in qualitative analysis (N=29).

Demographic characteristics	Adherent (n=15), n (%)	Nonadherent (n=14), n (%)	Total (N=29), n (%)
Disease diagnosis			
Crohn	7 (47)	9 (64)	16 (55)
Ulcerative colitis	7 (47)	5 (36)	12 (41)
Indeterminate colitis	1 (7)	0 (0)	1 (3)
Female	7 (47)	6 (43)	13 (45)
Race/ethnicity			
Non-Hispanic white	14 (93)	14 (100)	28 (97)
African American	1 (7)	—[a]	1 (3)
Protocol			
Weekly SMS[b] text messages	2 (13)	7 (50)	9 (31)
Every other week SMS text messages	13 (87)	7 (50)	20 (69)

[a]Not applicable.

[b]SMS: short message service.

Patterns/Themes

The findings include participants' perceptions of the 3 main components of the TELE-IBD system: (1) self-assessment questions, (2) action plans, and (3) educational messages. The main themes identified by participants are summarized in Table 2. We also addressed specific questions posed by the University of Maryland School of Medicine TELE-IBD team (RC, KR, CQ) regarding text-based platform, depression and adherence, TELE-IBD system in place of office visit, TELE-IBD perceptions, and participants' recommendations for future TELE-IBD systems.

Table 2. Summary of themes from qualitative interview of patients from the telemedicine for patients with inflammatory bowel disease trial.

Theme	Strengths	Weaknesses
TELE-IBD[a] self-assessment questions	• Improved awareness of symptoms and health status • Two-way connection with the treatment team • Creates digital log of symptoms and weight • Easy to respond	• Repetitive when asymptomatic • Unwanted or ill-timed reminder of disease • Relevance when experiencing atypical symptoms • Source of alarm regarding potential symptoms • Risk of fixating on weight
TELE-IBD action plans	• Reminder to maintain routine, take medications • Prompted calls from the treatment team	• Accuracy of zone rating—unable to provide context for responses or correct data entry errors • Unable to indicate call back is not needed
TELE-IBD educational messages	• Improved health literacy: (1) understanding of the disease process and treatment and (2) vocabulary for communicating with treatment team • Provides information to share across social network • Similar to a trivia game	• Repetitive over time • Not tailored to specific disease type • Not helpful for those already well-informed
Text platform	• Convenient • Ease of connection with provider	• Response window/timed lockouts • Cell phone provider problems • Inconvenient timing of messages • No opportunities to elaborate or correct responses

[a]TELE-IBD: telemedicine for patients with inflammatory bowel disease.

Self-Assessment Questions

Both adherent and nonadherent participants had positive and negative perceptions of the core assessment questions. In terms of the benefits of the core assessment questions, participants suggest the assessments improved their awareness of their symptoms and health status. Numerous examples of this perception are found across adherent and nonadherent participants:

> *It was a beneficial kind of reminder to, you know, like "hey, let me think about this - Ah, I'm feeling good."* [Nonadherent participant]

XSL•FO
RenderX

> *I actually kind of liked having to do it because it was kind of like, you know, bam another good week, bam another good week. It was almost like setting personal records or something.* [Nonadherent participant]

> *All the questions that they asked seemed like they actually pertained to things that all I had, I could possibly have going on, and I might actually say "yes" to a lot of them...it would kind of give me a little heads up I guess if I should definitely go ahead and schedule an appointment to get in there to see them.* [Adherent participant]

> *It inspired me to...keep an eye on my weight and so really pay attention to what was going on and not wait until I was in fold over pain.* [Adherent participant]

Participants also reported that the assessment questions created and sustained an important 2-way connection with their health care provider. Through remote monitoring, the patient had an efficient, ongoing means for alerting the provider to problems. Simultaneously, the provider could alert the patient if the symptoms were cause for concern. Knowing that the provider's office was receiving and reviewing participants' responses was reassuring. A total of 1 participant described there was peace of mind in "monitoration." Participants also appreciated that their responses formed a log that the provider could then review and discuss. Flare-ups and weight changes that occurred between scheduled office visits were documented, providing a more comprehensive portrait of the participant's disease experience.

Participants across the sample viewed the assessment questions as straight-forward and easy to answer. Participants did not report embarrassment in recording their responses. The general consensus across the sample was that weekly assessments worked well, but as is discussed next, entering regular assessments when one is feeling well can be tedious.

Critiques of the assessment questions often revolved around perceptions that the questions were (1) repetitive, particularly when not in a flare or (2) "generic," that is, did not address symptoms that the participant experienced. Both adherent and nonadherent participants raised these concerns. For example, regarding the tedious nature of the assessments, adherent and nonadherent participants described:

> *When you're not in a flare and these messages keep coming and it's the same thing, and you're like, 'oh, God, I'm answering the exact same questions over and over and over again. That's when, you know, like after like tenth time, you're just like, 'gee.'* [Adherent participant]

> *Sometimes I felt like I got too many texts, I was like, "I'm okay!" you know. But I would reply.* [Nonadherent participant]

The assessment questions also at times could be "scary" if the participant had not encountered a specific symptom (eg, questions about fistulas) or served as unwanted reminders of one's disease during a period of remission.

Many believed the content of questions were appropriate. However, for participants who had unusual symptoms or other autoimmune disorders, the questions did not seem relevant. These participants indicated they would have preferred more individualized assessment questions.

Participants' conflicting attitudes toward tracking their weight also should be recognized. Although many participants perceived regular weigh-ins as a helpful, independent indication of their disease status, others were concerned about fixating on the scale:

> *It's one of those things where I think it was helpful and I think it's a really good thing for the doctor to see but sometimes I don't like weighing myself. Right now I don't keep a scale in my house because I don't really necessarily want to know the number. I can tell when I'm losing weight and I don't really like obsessing about that number. I know for medical purposes you're going to want to track things that have a numerical value but I mean I can tell by the way I feel and how my clothes are fitting and how I look if I've been losing weight or anything...* [Adherent participant]

Finally, even if participants critiqued the core assessment questions, they perceived that the assessment process could be helpful for others:

> *I'm going to be on top of it whether I have this text messaging or not. [...] It's not going to help all. Some of them, it's not going to matter, but I believe it will help some of them. It will push some of them along, you know, to go and get the proper treatment and not delay, and to be honest with your doctor about your symptoms and everything.* [Adherent participant]

Action Plans

Participants had difficulty recollecting the action plans, often needing reminders from interviewers that the system provided a green, yellow, or red zone indicator at the end of the assessment. Although appreciative of the reminder to take their medication or to continue with their routines, participants who were mainly in the green zone did not remember receiving an action plan.

Those who did recollect receiving yellow or red indicators when they were not feeling well, which resulted in a call-back from the providers' nurse staff, were often appreciative of the ease and immediacy of the contact with the provider's office. As 1 adherent participant described, knowing symptoms would produce a TELE-IBD action plan was helpful:

> *I knew, OK if it was Wednesday, I know I'm getting the call Thursday so I would just wait and do that instead of calling the nurse and saying, "Hey this is what's going on." If I knew it was something that would put me in the yellow zone, I would just wait...* [Adherent participant]

> *I really liked it believe it or not because I was able to provide feedback on a weekly basis and I was often, depending on my answers, getting calls immediately practically from the doctor or nurse. She's like, "what's going on?"...OK, I'll talk to the doctor and*

XSL•FO
RenderX

> *we'll do whatever if need be. I thought that was awesome.* [Nonadherent participant]

The primary critique of the action plans was that the zone ratings were not perceived as accurate, and participants felt they had no means to provide explanatory information. Participants indicated, for example, that at times their assessments and subsequent zone ratings reflected other illnesses, life events, or even data entry errors. Rather than receiving the call from the provider's office, participants would have liked to have been able to indicate why no follow-up was needed.

Educational Messages

Participants described mixed perceptions of the educational messages. Enthusiasm for the educational messages did not vary by adherence status or whether the participant was in remission at baseline.

Participants who liked the educational messages appreciated learning more about the disease, the intestinal system, diet strategies, and terminology that could then be used when talking with their health care provider. Participants who liked the messages described them as "enlightening," "kinda fun," and helpful "trivia." To these participants, the educational messages reinforced their health literacy; as one described:

> *I got all these things in the back of my mind that hey, this could be something, you know, it just makes it easier to, you know, catch that in the future.* [Adherent participant]

Another indicated the messages served as important refreshers, or "a ha" moments that helped her more quickly identify symptoms:

> *I was getting a lot of little snippets of information, parts of information I already knew but maybe forgot about so when I got the information it was like oh yeah, oh that's right. So it was helpful in that sense and it helped me, I feel like I'm in more control of my body. I know when I'm getting the symptoms. I know when something starts to irritate me a lot quicker than I did in the past.* [Adherent participant]

Disease information then traveled along participants' social networks, as participants indicated they would share messages with spouses or other family members with similar conditions.

The most common negative comment on educational messages was that over time the messages became repetitive. Participants recommended reducing the repetition and adding more information on vitamins, probiotics, or diet. Others commented that the messages were not relevant to their disease. For participants who had a long history with their illness or a strong understanding of its triggers, the education messages did not provide new insights, such as this nonadherent participant:

> *I felt that they were kind of generic, but I also had been diagnosed and had been dealing with the UC for, you know, probably over five years at that point.*

Participants who worked in an allied health field also felt the messages were not informative. It is important to note that participants who critiqued the educational messages commonly indicated that they believed the messages could be of great value to others. As a nonadherent participant with a clinical background suggested, "I can understand how it would be helpful for someone who was coming at it from a different perspective."

Platform: Text-Based System

Participants generally approved of the mobile phone platform for the text-based telemedicine system. Among the participants for whom texting is a part of daily life, the system was a "short" and "convenient" way to monitor their symptoms and stay connected to their provider. Participants who did not like the mobile system objected to specific design features or noted problems attributed to their carrier.

Regarding design features, the chief design criticism was that data entry errors could not be corrected, and there was no means to communicate errors to prevent a call-back from their provider. Participants also perceived that the numerical response system did not fully capture their health status and were frustrated that there was not space in the system to elaborate on their current situation. Participants further noted that the timing of the SMS text messages became inconvenient over the course of the study if their work or family responsibilities changed.

Technical issues such as delayed responses from the TELE-IBD system were described as "extremely irritating," significantly extending the amount of time required to complete the assessment. Participants perceived the response window as too short to accommodate the vagaries of life and issues caused by their cell phone providers. As 1 adherent participant described:

> *It wasn't necessarily inconvenient, but there were times that I would see that I got the messages and it was like a time sensitive, if you didn't respond, you know, it wouldn't prompt the next questions. So there were times that I missed like that particular like set of questions just because it, you know, it like timed out or I wasn't able to finish answering because I was in class or busy or doing whatever.*

Participants who were not satisfied with the delivery mechanism noted that having an option for email delivery or some other type of survey delivery would be helpful. However, this should not be interpreted as a recommendation to switch entirely to an email system as others felt strongly that current delivery method was effective.

Depression or Mood Effect

The UM-SOM TELE-IBD team (RC, KR, CQ) found that depression correlates with adherence, especially among younger participants [21]. One 20-year-old participant (nonadherent) is a good example of this finding. He perceived the messages as a lifeline. He said he struggled to find people to talk to who would understand what he was going through:

> *When I would get these messages, it made me feel better because it made me realize that I'm not actually the only one that had these kinds of problems. So, when I would get them, I would get a bit excited, a good bit excited and I would feel relieved. Like a ten-pound weight had been lifted off my shoulders or off my chest.*

XSL•FO
RenderX

In this study, emotions were tied closely to one's health status. When in a flare, participants reported feeling stressed, depressed, or anxious. These participants were more likely than those who were not in a flare to be adherent to the TELE-IBD system core self-assessment questions. An adherent participant described what it was like to be in a flare, addressing indirectly how responding to the TELE-IBD system may have been part of the internal focus on one's disease:

> *Well when I was in the flare-up, it was, I'm more, what's the word, introverted, you don't talk as much, you're very concerned about what's going on in your body, so you don't focus on much more than that.*

A nonadherent participant likewise eventually reasoned that feeling ill might lead one toward being more responsive:

> *I do suppose if, if someone were really feeling down about it, it could feel like one more burden of oh goodness why am I doing this. But on the other hand, if they're having issues like that then that's all the more reason to be in communication with their physician.*

We also asked participants to reflect on the finding regarding depression. Several expressed surprise that being depressed is associated with adherence. A nonadherent participant responded by saying, "Ah!" and then, "Yeah, well it's, it does feel like, you know, someone's asking you how you're doing and it's a way to engage you for this from afar so I can definitely see that." Another nonadherent participant could see how this might be: "I think when people feel poorly, they're probably happy to answer as many questions as it's going to take for them to get helpful information." At the same time, an adherent participant cautioned that receiving TELE-IBD SMS text messages while being depressed might make someone more depressed:

> *If somebody was more depressed, I definitely think it would have affected them differently but I'm not sure if would be positive or negative. Maybe more negative because it's a constant reminder that you kind of have the disease.*

Finally, participants also indicated that the emotional impact of the disease was not well-incorporated into the system, for example, the core self-assessment questions. One participant explained:

> *I would say, and this isn't just for the study but in general, I think a bigger focus on emotional impact of the disease and also the fatigue. I think a lot times you focus on, you know, weight and, and pain levels and other things but a lot of it is just, it's very fatiguing and, you know, there's a lot of malnutrition, malabsorption that goes on. So I think fatigue is just one, is, you know, it's one of the more debilitating aspects and it's surprisingly so, I think people focus on, it must have hurt, you know, stomach pain all the time but you kind of build a tolerance to that but it's hard to build a tolerance to being tired all the time.* [Adherent participant]

Impact on Face-to-Face Office Visits or Other Forms of Clinical Care

Participants had mixed perceptions on how participating in a TELE-IBD system would impact their routine health care. Some indicated that if they were not having any IBD-related issues, checking in through the TELE-IBD system could substitute for a face-to-face visit. This was an especially attractive concept for participants who travel a long distance for provider office visits. However, others expressed concern that they would still want to meet with their provider twice a year or in the event of a flare. For example, 1 adherent participant who lived an hour from her provider expressed the system could help monitor, but:

> *I would still want to go in and make sure my medication is OK...Over the six months a lot of times I might develop like a question that I might want to ask and through that, you know, the text message you're not able to necessarily do that. So, it's nice to be able to go in and have that dialogue about certain things that I might be experiencing.*

Adherence, Health Status, and Telemedicine for Patients With Inflammatory Bowel Disease Perceptions

A few patterns emerged regarding adherence, participants' perceptions of the relevance of the system, and experiences with design and technical glitches.

First, regarding the relevance of the system, nonadherent participants were much more assertive in indicating they actively managed their disease and/or had a strong understanding of the disease (whether because of their own research, the length of time diagnosed, or their professional background), such that the system was less helpful to them than it may be to others.

Second, although both adherent and nonadherent participants experienced technical difficulties, discussions of the technical problems were much more salient among the nonadherent participants. The system became "extremely irritating" when the SMS text message questions timed out too quickly, extending the time required to complete the questions to 15 min or more.

Regarding the role of health status, both the adherent and nonadherent participants noted the value of the TELE-IBD system during a flare or when changing medications. The TELE-IBD system was clearly recognized as allowing for a more immediate and effortless way to communicate with the doctor's office. It also is very reassuring to receive the doctor's call back when experiencing symptoms. Although both groups reported the system was "repetitive," the nonadherent tended to voice more strongly that the regular pings from the system were "annoying" when they were feeling well.

Discussion

Principal Findings

To our knowledge, this is the first telemedicine for IBD post-RCT qualitative study of adherent and nonadherent intervention participants. The main finding was that patients

XSL•FO
RenderX

prefer a flexible follow-up system that is personalized, including education messages that can be targeted to individual patients and decreased repetition. A telemedicine system that is 2-way communication, easily connecting patients with providers and provides "actionable" responses to patient self-assessment is preferred. Telemedicine implemented as an adjunct to patient-initiated consultations and self-management and regular office visits was identified as optimal approaches to meet the patients' needs. The TELE-IBD system was reported by participants as effective in facilitating IBD self-management. Participants also considered the TELE-IBD useful in treatment decisions made with providers in-between office visits. New models of IBD telemedicine care could improve the patients' experience of care and reduce or delay unnecessary health care utilization.

The TELE-IBD intervention in this study followed a participatory care framework by testing pilot feasibility [22,23] and acceptability, followed by a randomized clinical trial [13,15] and follow-up qualitative evaluation of adherence and nonadherence to the text-based intervention [24]. Evaluating performance at each of these phases is critical to the overall success of the intervention and to ensure that the system is safe and beneficial to patients. Although adherence in this study was targeted to 80% (actual=60%), this qualitative assessment demonstrates ongoing patient engagement to help mitigate problems associated with risk of attrition when evaluating new models of care such as telemedicine.

Limitations

We recognize the reported study has limitations. The qualitative assessments were from patients willing to be interviewed and may not be representative of all patients in the TELE-IBD clinical trial or other IBD patients considering a telemedicine system of care. We attempted to address this by recruiting adherent and nonadherent persons to be interviewed. Interviews occurred after use of the system. Memory or recall may have affected participants' responses. Finally, it's important to note that although clinicians and scientists have expectations for adherence, not all patients want to be "engaged" all the time. Patient engagement is a continuous behavior, and our assessment of adherence may have captured patient perceptions at a single point in time.

Conclusions

Participants, both adherent and nonadherent, identified clear benefits to the TELE-IBD system, including obtaining a better understanding of the disease process, monitoring their symptoms, and feeling connected to their health care provider. In this sense, TELE-IBD facilitates access to care, particularly for those who live a long distance from the provider and/or are reluctant to or have difficulty contacting the provider between scheduled office visits. Both educational messages and core assessment questions contribute to these perceptions of the TELE-IBD system. Such a connection is particularly welcome during an IBD flare and not as necessary when asymptomatic.

Participants experienced considerable recall issues around the action plans, although could remember the color zones once prompted and appreciated knowing that they would receive a provider call if experiencing critical symptoms. Reminders regarding medication were appreciated by some, but others said the reminders were not necessary. The study reported here demonstrates the importance of qualitative assessment of participants' views in technology-based interventions. On the basis of our results, future studies of revised remote monitoring systems in patients with IBD are needed. To promote patient engagement, future systems should include flexible options for testing (mode of testing and frequency of testing), reduce "irritant" factors such as timed lock outs and set testing schedules, and allow for corrections in responses. Furthermore, assessment of symptoms and receipt of action plans and educational messages should be as personalized as possible. Finally, remote monitoring should remain an adjunct to in-person monitoring.

Acknowledgments

This study was supported by grants from Agency for Healthcare Research and Quality (1R01HS018975-01A1) and the Broad Medical Research Program at the Crohn's & Colitis Foundation of America (CCFA; 431257).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study interview guide.

[DOCX File, 14KB - jmir_v21i6e14165_app1.docx]

References

1. Loftus EV. Clinical epidemiology of inflammatory bowel disease: incidence, prevalence, and environmental influences. Gastroenterology 2004 May;126(6):1504-1517. [Medline: 15168363]
2. Sands BE. From symptom to diagnosis: clinical distinctions among various forms of intestinal inflammation. Gastroenterology 2004 May;126(6):1518-1532. [Medline: 15168364]

XSL•FO
RenderX

3. Bernklev T, Jahnsen J, Aadland E, Sauar J, Schulz T, Lygren I, IBSEN Study Group. Health-related quality of life in patients with inflammatory bowel disease five years after the initial diagnosis. Scand J Gastroenterol 2004 Apr;39(4):365-373. [Medline: 15125469]
4. Limsrivilai J, Stidham R, Govani S, Waljee A, Huang W, Higgins P. Factors that predict high health care utilization and costs for patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol 2017 Dec;15(3):385-92.e2 [FREE Full text] [doi: 10.1016/j.cgh.2016.09.012] [Medline: 27645518]
5. Mehta F. Report: economic implications of inflammatory bowel disease and its management. Am J Manag Care 2016 Dec;22(3 Suppl):s51-s60 [FREE Full text] [Medline: 27269903]
6. Malarcher C, Wheaton A, Liu Y, Greenlund SF, Greenlund SJ, Lu H, et al. Hospitalizations for Crohn's Disease - United States, 2003-2013. MMWR Morb Mortal Wkly Rep 2017 Apr 14;66(14):377-381 [FREE Full text] [doi: 10.15585/mmwr.mm6614a1] [Medline: 28406887]
7. Kane S, Huo D, Aikens J, Hanauer S. Medication nonadherence and the outcomes of patients with quiescent ulcerative colitis. Am J Med 2003 Jan;114(1):39-43. [Medline: 12543288]
8. Cross R, Wilson K, Binion D. Polypharmacy and Crohn's disease. Aliment Pharmacol Ther 2005 May 15;21(10):1211-1216 [FREE Full text] [doi: 10.1111/j.1365-2036.2005.02429.x] [Medline: 15882241]
9. Cross R, Lapshin O, Finkelstein J. Patient subjective assessment of drug side effects in inflammatory bowel disease. J Clin Gastroenterol 2008 Mar;42(3):244-251. [doi: 10.1097/MCG.0b013e31802f19af] [Medline: 18223502]
10. Quan H, Present J, Sutherland L. Evaluation of educational programs in inflammatory bowel disease. Inflamm Bowel Dis 2003 Nov;9(6):356-362. [doi: 10.1097/00054725-200311000-00003] [Medline: 14671484]
11. Abutaleb A, Buchwald A, Chudy-Onwugaje K, Langenberg P, Regueiro M, Schwartz DA, et al. Inflammatory bowel disease telemedicine clinical trial: impact of educational text messages on disease-specific knowledge over 1 year. Inflamm Bowel Dis 2018 Dec 15;24(10):2191-2197. [doi: 10.1093/ibd/izy149] [Medline: 29788137]
12. Rubin D, Feld L, Goeppinger S, Margolese J, Rosh J, Rubin M, et al. The Crohn's and Colitis Foundation of America Survey of inflammatory bowel disease patient health care access. Inflamm Bowel Dis 2017 Dec;23(2):224-232. [doi: 10.1097/MIB.0000000000000994] [Medline: 27997434]
13. Cross R, Langenberg P, Regueiro M, Schwartz DA, Tracy JK, Collins JF, et al. A randomized controlled trial of TELEmedicine for Patients with Inflammatory Bowel Disease (TELE-IBD). Am J Gastroenterol 2019 Mar;114(3):472-482. [doi: 10.1038/s41395-018-0272-8] [Medline: 30410041]
14. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps M, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet 2017 Sep 2;390(10098):959-968. [doi: 10.1016/S0140-6736(17)31327-2] [Medline: 28716313]
15. Cross R, Jambaulikar G, Langenberg P, Tracy JK, Collins JF, Katz J, et al. TELEmedicine for Patients with Inflammatory Bowel Disease (TELE-IBD): design and implementation of randomized clinical trial. Contemp Clin Trials 2015 May;42:132-144. [doi: 10.1016/j.cct.2015.03.006] [Medline: 25812483]
16. Harvey RF, Bradshaw JM. A simple index of Crohn's-disease activity. Lancet 1980 Mar 8;1(8167):514. [Medline: 6102236]
17. Best W. Predicting the Crohn's disease activity index from the Harvey-Bradshaw Index. Inflamm Bowel Dis 2006 Apr;12(4):304-310. [doi: 10.1097/01.MIB.0000215091.77492.2a] [Medline: 16633052]
18. Walmsley R, Ayres R, Pounder R, Allan R. A simple clinical colitis activity index. Gut 1998 Jul;43(1):29-32 [FREE Full text] [doi: 10.1136/gut.43.1.29] [Medline: 9771402]
19. Marín-Jiménez I, Nos P, Domènech E, Riestra S, Gisbert JP, Calvet X, et al. Diagnostic performance of the simple clinical colitis activity index self-administered online at home by patients with ulcerative colitis: CRONICA-UC study. Am J Gastroenterol 2016 Feb;111(2):261-268. [doi: 10.1038/ajg.2015.403] [Medline: 26753886]
20. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
21. Chudy-Onwugaje K, Abutaleb A, Buchwald A, Langenberg P, Regueiro M, Schwartz DA, et al. Age modifies the association between depressive symptoms and adherence to self-testing with telemedicine in patients with inflammatory bowel disease. Inflamm Bowel Dis 2018 Dec 29;24(12):2648-2654. [doi: 10.1093/ibd/izy194] [Medline: 29846623]
22. Cross R, Finkelstein J. Feasibility and acceptance of a home telemanagement system in patients with inflammatory bowel disease: a 6-month pilot study. Dig Dis Sci 2007 Feb;52(2):357-364. [doi: 10.1007/s10620-006-9523-4] [Medline: 17211702]
23. Cross R, Arora M, Finkelstein J. Acceptance of telemanagement is high in patients with inflammatory bowel disease. J Clin Gastroenterol 2006 Mar;40(3):200-208. [Medline: 16633120]
24. Jackson BD, Gray K, Knowles SR, De Cruz P. EHealth technologies in inflammatory bowel disease: a systematic review. J Crohns Colitis 2016 Sep;10(9):1103-1121. [doi: 10.1093/ecco-jcc/jjw059] [Medline: 26928960]

Abbreviations

CD: Crohn disease
IBD: inflammatory bowel disease
IC: indeterminate colitis

XSL•FO
RenderX

QoL: quality of life
SMS: short message service
TELE-IBD: telemedicine for patients with inflammatory bowel disease
UC: ulcerative colitis

Edited by G Eysenbach; submitted 28.03.19; peer-reviewed by K VanRavenstein, N Pedersen, S Li; comments to author 23.04.19; revised version received 01.05.19; accepted 16.05.19; published 03.06.19.

Please cite as:
Quinn CC, Chard S, Roth EG, Eckert JK, Russman KM, Cross RK
The Telemedicine for Patients With Inflammatory Bowel Disease (TELE-IBD) Clinical Trial: Qualitative Assessment of Participants' Perceptions
J Med Internet Res 2019;21(6):e14165
URL: https://www.jmir.org/2019/6/e14165/
doi:10.2196/14165
PMID:31162128

©Charlene Connolly Quinn, Sarah Chard, Erin G Roth, J Kevin Eckert, Katharine M Russman, Raymond K Cross. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 03.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Optional Web-Based Videoconferencing Added to Office-Based Care for Women Receiving Psychotherapy During the Postpartum Period: Pilot Randomized Controlled Trial

Rebecca Yang[1], MPH; Simone N Vigod[1,2,3,4], MD, MSc; Jennifer M Hensel[1,2,3,5], MD, MSc

[1]Women's College Hospital Institute for Health System Solutions and Virtual Care, Women's College Hospital, Toronto, ON, Canada

[2]Department of Psychiatry, Women's College Hospital, Toronto, ON, Canada

[3]Women's College Research Institute, Women's College Hospital, Toronto, ON, Canada

[4]Institute for Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada

[5]Department of Psychiatry, University of Manitoba, Winnipeg, MB, Canada

Corresponding Author:
Rebecca Yang, MPH
Women's College Hospital Institute for Health System Solutions and Virtual Care
Women's College Hospital
76 Grenville Street, 6th Floor
Toronto, ON, M5S 1B2
Canada
Phone: 1 416 323 6400 ext 5989
Email: rebecca.yang@wchospital.ca

Abstract

Background: Depression and anxiety during the postpartum period are common, with psychotherapy often being the preferred method of treatment. However, psychological, physical, and social barriers prevent women from receiving appropriate and timely psychotherapy. The option of receiving psychotherapy through videoconferencing (VC) during the postpartum period presents an opportunity for more accessible and flexible care.

Objective: The aim of this study was to assess the feasibility, acceptability, and preliminary effectiveness of optional VC added to usual office-based psychotherapy, with a psychotherapist during the postpartum period.

Methods: We conducted a pilot randomized controlled trial with 1:1 randomization to office-based care (treatment as usual; TAU) or office-based care with the option of VC (treatment as usual plus videoconferencing; TAU-VC) for psychotherapy during the postpartum period. We assessed the ability to recruit and retain postpartum women into the study from an urban perinatal mental health program offering postpartum psychotherapy, and we evaluated the uptake, acceptability, and satisfaction with VC as an addition to in-person psychotherapy. We also compared therapy attendance using therapist logs and symptoms between treatment groups. Symptoms were assessed at baseline and 3 months postrandomization with the Edinburgh Postnatal Depression Scale, Generalized Anxiety Disorder 7-item, and Parental Stress Scale. Furthermore, 3-month scores were compared between groups with intention-to-treat linear mixed-effects models controlling for baseline score.

Results: We enrolled 38 participants into the study, with 19 participants in each treatment group. Attendance data were available for all participants, with follow-up symptom measures available for 25 out of 38 participants (66%). Among the 19 TAU-VC participants, 14 participants (74%) utilized VC at least once. Most participants were highly satisfied with the VC option, and they reported average savings of Can $26 and 2.5 hours in travel and childcare expenses and time per appointment. There were no significant differences between the 2 groups for psychotherapy attendance or symptoms.

Conclusions: The option of VC appears to be an acceptable method of receiving psychotherapy for postpartum women, with benefits described in costs and time savings. On the basis of this small pilot sample, there were no significant differences in outcomes between office-based care with or without the option of VC. This study has demonstrated the feasibility of such a program in an urban center, which suggests that a larger study would be beneficial to provide evidence that is more conclusive.

(J Med Internet Res 2019;21(6):e13172) doi:10.2196/13172

XSL•FO
RenderX

KEYWORDS

mental health; psychotherapy; postpartum period; videoconferencing

Introduction

The need for high-quality, accessible mental health care is significant, with important individual and societal implications. However, the lack of services, underawareness of services, and stigma and attitudes associated with mental health are large barriers to help seeking [1], with at least 50% of those affected not receiving adequate care [2]. The postpartum period can be a particularly vulnerable time for women requiring mental health support. Up to 15% of women experience postpartum mood disorders [3,4], with most preferring psychotherapy over medication for the treatment of their mental illnesses [5]. However, childcare responsibilities and physical recovery, particularly after a surgical birth, create barriers for regular therapy attendance [6-8]. Barriers to care often lead to delayed care, treatment discontinuation, and the lack of follow-up on recommendations. Untreated symptoms of perinatal mental illness can last for months or years, leading to persistent mental health problems and decreased quality of life for both the woman and child [9,10].

Virtual care tools, which facilitate remote interactions between patients and health care providers through information technology [11], have gained traction as potential solutions to deliver effective treatment for most mental health needs [12-14]. They can address many barriers associated with in-person care, such as travel distance and time, travel costs, and stigma [5,15]. In addition, patients may be more motivated to seek and continue treatment if they are in a familiar environment and can avoid potentially stressful situations, such as driving in urban cities and navigating a hospital [16]. More women are increasingly accessing the internet for support and information during the postpartum period [17]. A study by Maloni and colleagues [6] found that among postpartum women with depression and women who experienced pregnancy complications, over 90% of the women demonstrated an interest in internet-based interventions, with 40% indicating a specific interest in chatting virtually with a provider with expertise in postpartum depression. Similarly, Sawyer et al found that access to a nurse and peers through an internet-based chat forum yielded improved parental stress scores compared with standard treatment alone [18]. A survey conducted in our own ambulatory outpatient mental health program, which has a special focus on perinatal mental health, found that pregnancy factors and/or childcare were cited as a common barrier to office-based care [19]. Among all survey respondents with an identified barrier, 93% of these patients were interested in internet-based treatment tools, and 76% of these patients were specifically interested in videoconferencing (VC) to receive their care from home, using a personal device [19]. Studies on VC-based psychotherapy have demonstrated similar treatment outcomes when directly compared with the face-to-face option [16,20].

To our knowledge, no studies have specifically assessed the option of VC in addition to office-based care to determine if it added any benefits. By overcoming the barriers to care that can be present in varying ways across a course of treatment, the addition of VC to an office-based treatment program could provide a very patient-centered option to receive care. This could lead to improved therapy adherence and completion rates and possibly better outcomes as a result. In this study, we assessed the acceptability, uptake, and preliminary effectiveness of adding the option to receive psychotherapy delivered by VC to standard office-based psychotherapy for women experiencing mood and anxiety problems during their first year postpartum.

Methods

Study Design

This study was a pilot randomized controlled trial, comparing usual office-based care (treatment as usual; TAU) with office-based care with the option of VC on a personal device (treatment as usual plus videoconferencing; TAU-VC) for psychotherapy for mood or anxiety problems in the postpartum period. Participants were randomized 1:1 into 1 of the 2 treatment arms, and outcomes were measured at baseline and 3 months postrandomization. This study had 3 objectives: (1) assess the ability to recruit and retain postpartum women into the study, (2) evaluate the uptake and satisfaction with VC as an addition to in-person psychotherapy, and (3) compare therapy attendance and symptoms after 3 months between those receiving the option of VC with those receiving usual care. Research ethics approval was obtained from the Women's College Hospital Research Ethics Board.

Study Setting

Participants were recruited from a specialized mental health program in an ambulatory hospital within an urban city in Ontario, Canada, which provides publicly funded individual psychotherapy for mood, anxiety, obsessive-compulsive, or trauma or stressed-related disorders in pregnancy and up to 1 year postpartum delivered by highly trained psychotherapists with a master's degree in social work. The program accepts referrals from primary care, midwifery, obstetricians, and psychiatrists from across Ontario, although the vast majority of women who receive care reside in the greater Toronto region (population ~6.4 million). The program receives approximately 1200 referrals annually, and it will provide assessment and treatment to women within the first year postpartum. The clinic does provide prebooked childcare during daytime hours. Individuals referred to the program are assessed by a psychiatrist, and if deemed suitable for psychotherapy, they may be referred to group or individual services. A psychotherapist contacts referred individuals to schedule an initial appointment, usually within 1 month of their psychiatric assessment. Therapists receive about 10 referrals for individual treatment a month, of which, approximately 70% convert into a course of treatment.

Recruitment

Study recruitment took place from June 2016 to July 2017. We aimed to recruit 40 participants, as has been recommended for pilot studies of an intervention where the goal is to have

XSL•FO
RenderX

sufficient variability to examine processes in the implementation of the protocol [21]. Participation was initially just offered to only new referrals to the clinic, but it was expanded to include patients already in treatment after trial commencement to increase recruitment rate. Psychotherapists introduced the study to newly referred patients during the initial call about treatment and to existing patients during a therapy session. The inclusion criteria for this study were patients in the program, aged 18 years or older, referred to psychotherapy for mood and/or anxiety symptoms, who had access to and ability to use a Web-enabled personal device or computer with the required audiovisual communication capability, and had a functioning email address (a requirement for the VC platform). Patients could be approached when pregnant if they intended to continue or start psychotherapy postpartum, but the study protocol was not initiated until after the woman returned for therapy postpartum. For postpartum patients, only women less than 9 months postpartum were included, to allow a minimum of 3 months of treatment before they reached 1 year postpartum and were discharged from the program. Patients with acute mania or psychosis or severe suicidal ideation with planning and intent were excluded. Therapists provided the contact information of interested patients to a research assistant. The research assistant contacted interested patients to introduce and explain the study. If the patient was still interested, the research assistant obtained and documented informed verbal consent over the phone and emailed a copy of the completed consent form to the patient.

Intervention Groups

Treatment as Usual

In the TAU group, participants received the standard form of psychotherapy available in the clinic. In the program, psychotherapists provide first-line evidence-based treatments for depression and anxiety, including cognitive behavior therapy and interpersonal therapy, adapted for the postpartum context. This treatment is mostly provided in person, although sessions are sometimes delivered by telephone on an ad hoc basis if a patient is unable to attend in person because of medical or last-minute childcare problems. Patients may also receive psychiatric care, including medication management, from the program psychiatrist until approximately 1 year postpartum. At this point, patients are referred back to their primary care provider for ongoing management.

Treatment as Usual Plus Videoconferencing

In the TAU-VC group, participants had access to TAU with the added option of having any of their treatment sessions over VC. The VC platform was hosted by the Ontario Telemedicine Network, a government-funded organization that supplies telemedicine services free of charge to health care organizations in Ontario. The secure platform is available across the province and can be used for communication among health professionals, between patients and health professionals, or as a webcast platform. To access the Web-based VC platform, participants had to download a plug-in for their computer, which required a webcam and microphone or an app for their mobile phone or tablet. All the therapists had unique accounts to log into the portal on their office computers. Therapists were able to schedule a therapy session and invite the participants by emailing them a personal link that included the date and time. Participants could attend sessions from their desired location, but they were encouraged to ensure the location was private enough so that they could fully participate in the session. When the appointment time arrived, participants and therapists could enter the virtual session. In addition to video and audio sharing, therapists had the ability to share their screens, which allowed them to show worksheets or other visual materials during the session. All participants who received access to VC received an instruction pamphlet by email and had a brief test call with the research assistant to ensure their device met the required technical specifications, to give a demonstration of the platform, and to troubleshoot any initial technical issues. During the study, participants in the TAU-VC group were encouraged, but not required, to use VC for their therapy sessions, and they were still able to access in-person therapy sessions and phone support as per standard care. Participants were also informed that therapists could, at their discretion, recommend an in-person visit if they felt it was warranted for clinical reasons. Technical support could be obtained from the research assistant or from the Ontario Telemedicine Network. All program psychotherapists (3 individuals) provided VC sessions during the study.

Randomization and Blinding

Participants were allocated 1:1 to the intervention and control groups using simple randomization. For randomization, 20 slips of paper labeled with *TAU* and 20 with *TAU-VC* were placed in opaque envelopes by a research staff member who was not involved with this study. Upon documentation of informed consent, the research assistant opened 1 of the envelopes, revealed the group allocation, and communicated it to the participant and psychotherapist. If a participant withdrew before being informed of the participant's allocation, the paper was returned to an envelope and put back in the pile. Participants and therapists were unblinded; data analyses were blinded to group allocation.

Data Collection

Baseline data were collected for all participants by the research assistant. If the participant was postpartum, data were collected at the time of recruitment, and if the participant was pregnant when recruited, data were collected following delivery of the child and upon reengaging in therapy. Baseline data included participant age, age of child or children, childcare support, and exposure to previous therapy, which referred to whether the participant had any previous therapy experience in her lifetime. Additional data, including address (to calculate travel distance from the clinic) and whether the participant was a newly referred or existing patient at the time of recruitment, were provided by therapists or extracted from individual clinical charts.

Baseline symptom and function measures were administered by a Web-based survey. Participants provided an email address, and they were sent a link to a survey hosted on FluidSurveys, which was later transitioned to SurveyMonkey after an institutional change in survey host. All participants received another survey link by email as a reminder to complete follow-up measures 2 weeks before reaching 3 months, and they received the link again at 3 months if the survey had not been

XSL•FO
RenderX

completed. A follow-up phone call was made to any outstanding surveys at 2 weeks after the 3-month time point. At follow-up, participants in the TAU-VC group were asked additional questions about their use of and satisfaction with VC.

The psychotherapists completed a therapy log for each of their participants, which documented the length of time (in minutes) and format (in person, VC, or telephone) of each session, along with cancellations and no-shows and how often a child was present for the session.

Measures

Edinburgh Postnatal Depression Scale

The Edinburgh Postnatal Depression Scale (EPDS) is a self-report depression screening measure that has been validated for use in pregnancy [22]. EPDS scores >12 are predictive of a diagnosis of major depressive disorder. The EPDS is able to detect women with depression in the perinatal period better than traditional depression measures because of the increased weight given to anxiety symptoms that appear to be more common in perinatal than in nonperinatal depression [22].

Generalized Anxiety Disorder 7-Item

The Generalized Anxiety Disorder 7-item (GAD-7) assesses symptoms of general anxiety. It has 7 items rated on a 4-point scale from "never" to "nearly every day," as well as 1 perceived impairment rating. A score of ≥10 is highly suggestive of a problem with anxiety. A reduction of 5 points corresponds to a clinically meaningful improvement, with a reduction in score of 50% representing response [23].

Parental Stress Scale

The Parental Stress Scale (PSS) is an 18-item questionnaire that was developed to measure the level of stress associated with raising children. All items are scored on a 5-point Likert scale from "strongly disagree" to "strongly agree." The questionnaire has shown good reliability and internal consistency [24]. It has been used to study parenting stress in a variety of parent-child circumstances, and it is correlated with measures of depression [25]. A total score is obtained by summing all items.

Telemedicine Satisfaction Questionnaire

The Telemedicine Satisfaction Questionnaire (TSQ) was originally developed and studied by Yip et al [26] to assess satisfaction with telemedicine. It has 15 items rated on a 5-point Likert scale from "strongly disagree" to "strongly agree." The authors reported good internal validity and consistency. A factor analysis yielded 3 components: quality of care provided [8 items], similarity to in-person face-to-face interaction (5 items), and perception of the interaction (1 item). This study used the original TSQ with the term "telemedicine" substituted with "video visits." An overall score can be created by summing all items, or subscale scores can be calculated for the component domains. Individual items may also be examined to assess the positive and negative aspects of the intervention [26]. In this study, 2 additional items were added: "I am going to miss the equipment when the project ends," and "I would be willing to pay for video visits privately." These latter questions were added to gather additional data regarding interest in the intervention.

Patient Reported Costs Questionnaire

The Patient Reported Costs Questionnaire (PRCQ) was developed by the research team to assess the participant costs and cost savings associated with access to VC. Respondents were asked to estimate the amount and source of time and money they saved when they attended a session of therapy with VC compared with in person.

Data Analysis

Objective 1: Recruitment and Retention

We documented the number and rate of referrals, baseline demographics and symptoms scores, and the proportion of participants who provided follow-up data at the outcome time points.

Objective 2: Uptake and Satisfaction With Videoconferencing

Descriptive statistics were generated to examine uptake of the intervention as defined as number and percentage of sessions attended by VC. TSQ and PRCQ items were analyzed descriptively.

Objective 3: Comparison of Therapy Attendance and Symptoms

Session attendance variables were summarized as group means or medians and compared between groups with unpaired 2-tailed *t* tests or nonparametric tests where medians were reported. Endpoint EPDS, GAD-7, and PSS total scores at 3 months were separately compared between groups with intention-to-treat linear mixed-effects models, including the relevant baseline score as a covariate. We conducted the same analyses in the subgroup of participants with baseline EPDS scores greater than 12 to examine the participants with baseline symptoms suggestive of a major depressive disorder [22].

Results

Objective 1: Recruitment and Retention

Participant flow through the study is outlined in Figure 1. Between June 2016 and July 2017, a total of 43 individuals were referred to the study. Although all the participants provided consent to participate, 5 did not complete baseline measures; therefore, they were not allocated to a treatment condition. A total of 38 participants were therefore included in the study, 19 in each group. At 3 months postrandomization, therapy attendance data were available for all participants (n=38/38, 100%), and 25 out of 38 (66%) participants completed symptom scale scores, with a slightly higher completion rate in the TAU-VC (n=14/38, 78%) than the TAU (n=11/38, 61%) condition.

Table 1 summarizes the baseline data for individuals randomized by treatment group. The average baseline EPDS, GAD-7, and PSS scores in both groups indicate moderate postnatal depression, generalized anxiety, and parental stress, respectively.

XSL•FO
RenderX

Figure 1. Participant flow for the pilot randomized controlled trial. Refer to Study Design for the study objectives. TAU: treatment as usual; TAU-VC: treatment as usual plus videoconferencing.

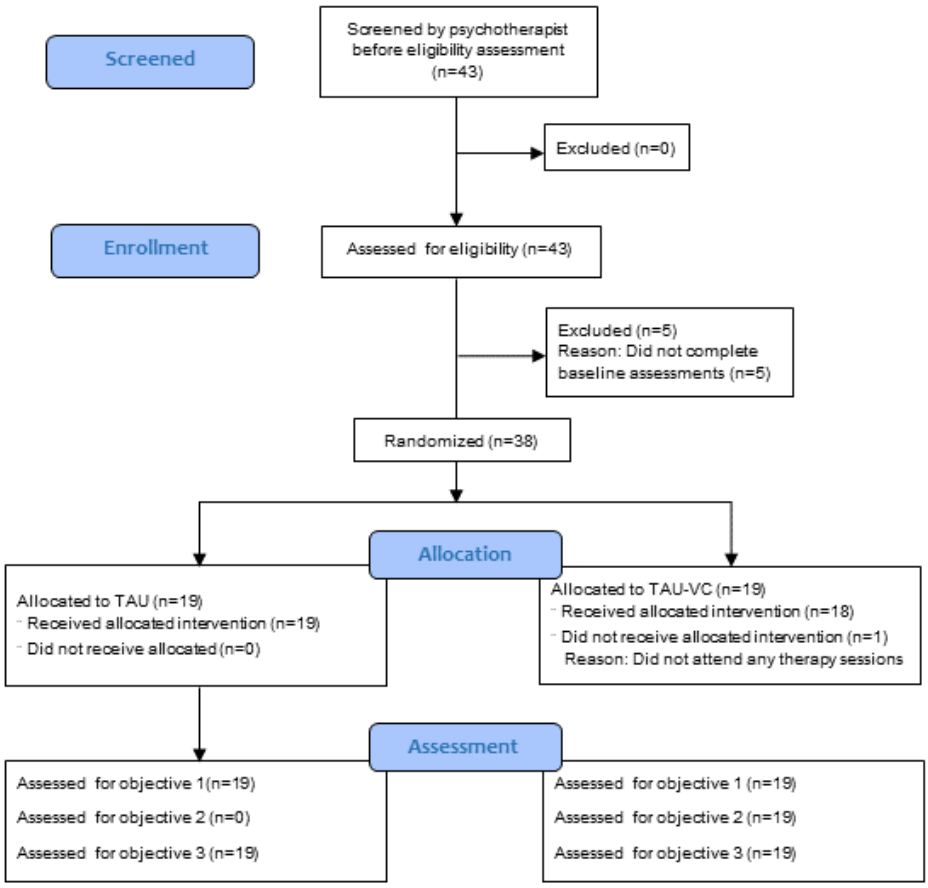


Table 1. Baseline characteristics of participants.

Variable	TAU-VC[a] (n=19)	TAU[b] (n=19)
Age (years), mean (SD)	33.8 (3)	33.8 (3)
Distance from clinic (in km), mean (SD)	7.8 (6)	7.7 (4)
Married or common-law, n (%)	18 (95)	16 (84)
Single, n (%)	1 (5)	3 (16)
Family or formal childcare support, n (%)	10 (53)	15 (79)
Index child age at start of study in weeks, median (IQR[c])	8.0 (14)	9.0 (21)
Any other children, n (%)	10 (49)	10 (51)
Previous therapy experience, n (%)	16 (89)	15 (79)
New patient, n (%)	9 (47)	6 (32)
Existing patient, n (%)	10 (53)	13 (68)
Baseline Edinburgh Postnatal Depression Scale, mean (SD)	13.3 (6)	13.6 (5)
Baseline Generalized Anxiety Disorder 7-item, mean (SD)	10.5 (5)	9.0 (5)
Baseline Parental Stress Scale, mean (SD)	45.3 (12)	49.5 (12)

[a]TAU-VC: treatment as usual plus videoconferencing.

[b]TAU: treatment as usual.

[c]IQR: interquartile range.

XSL•FO
RenderX

Objective 2: Uptake and Satisfaction With Videoconferencing

VC was used at least once by 14 (74%) of the participants in the TAU-VC group. The median number of VC sessions was 2 (range: 0-8). On average, users attended 50% of their sessions with VC, with 4 participants using VC for 100% of their therapy sessions and 4 participants using it for none of their sessions. Figure 2 displays the proportion of sessions completed by VC, phone, and in person among participants in the TAU-VC group. Although 19 participants were allocated to the TAU-VC condition, 1 participant did not attend any sessions of any kind and is not shown in the figure.

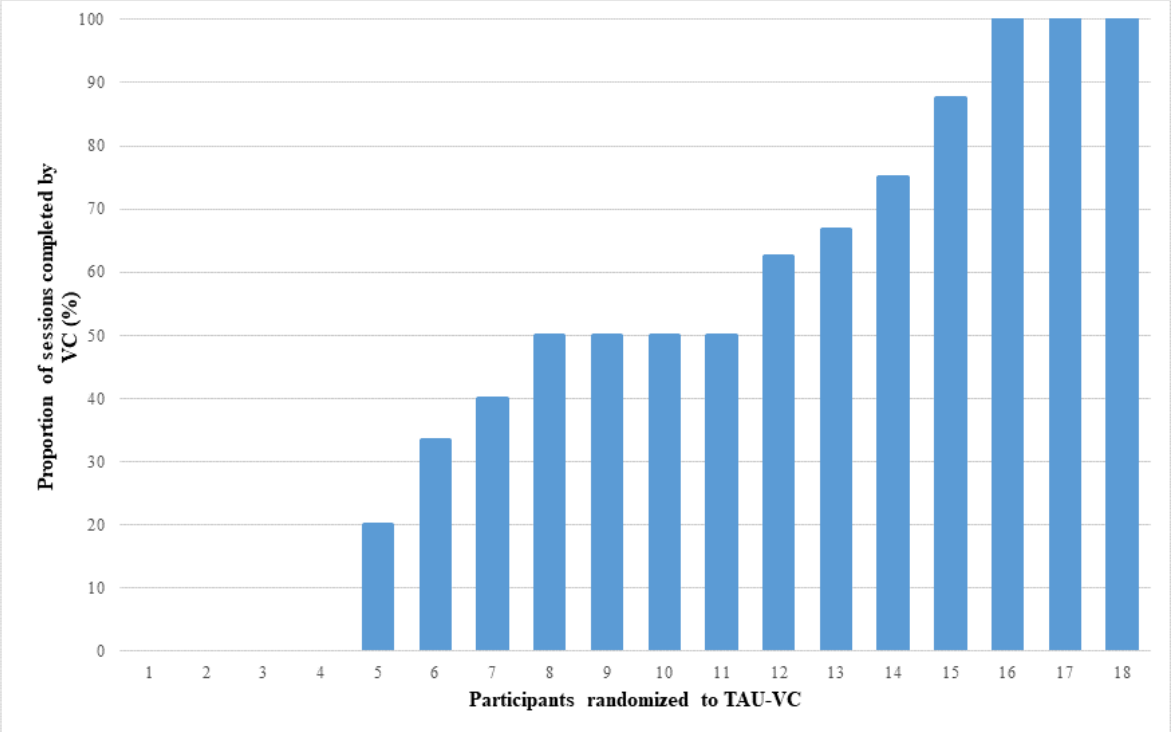
A total of 11 of the 14 VC users completed the TSQ. Of them, 9 VC users (82%) agreed or strongly agreed that they felt VC allowed them to attend therapy sessions more frequently. In a subgroup comparison between TAU-VC participants who were TSQ responders and TAU participants, TSQ responders attended more sessions face to face (4.6 vs 2.9, *P*=.06), with total therapist contacts being more similar (5.1 vs 4.8, *P*=.80), indicating that the TAU groups had a higher number of phone contacts. The average TSQ item score was 4.7 out of 5, with average scores of 4.7 (SD 0.43), 4.7 (SD 0.31), and 4.6 (SD 0.67) in the domains of quality of care provided, similarity to in-person face-to-face interaction, and perception of the interaction, respectively. All participants said they would miss having access to VC when the project ended, but only 1 participant expressed a high willingness to pay out of pocket to have it as an available option.

All participants who used VC at least once reported saving money and time when attending therapy sessions with VC compared with in person. Specifically, the average cost savings was Can $26 (SD 19.7, range: Can $7-$70) in travel, childcare, and other expenses per session, and the average time savings was 2.5 hours (SD 1.6, range: 1-7 hours) in preparation and transportation per session. There was a small start-up cost to purchase the webcams for the therapist computers (approximately Can $200 per therapist). There were no operating costs for the VC platform, as it was provided courtesy of a government-funded organization in Ontario.

Objective 3: Comparison of Therapy Attendance and Symptoms

There were no significant differences between the 2 groups for the total number of sessions attended, the average length of time per session, the number of no-shows or cancellations, and the number of sessions with a child in attendance (Table 2). Participants in the TAU-VC group tended to attend more face-to-face sessions (in person or by VC) and fewer phone sessions than the TAU participants, but the differences did not reach statistical significance.

Figure 2. Proportion of sessions completed by videoconferencing among participants who were randomized to treatment as usual plus videoconferencing group. TAU-VC: treatment as usual plus videoconferencing; VC: videoconferencing.



XSL•FO
RenderX

Table 2. Summary of therapy attendance variables between groups.

Variable[a]	TAU-VC[b] (n=19)	TAU[c] (n=19)	*P* value
Total number of sessions (all types), mean (SD)	4.5 (2.1)	4.8 (2.9)	.71
Total time (min)/session, median (IQR[d])	55.7 (10.0)	60.0 (12.5)	.58
Cancellations/no-shows, median (IQR)	1.0 (2.0)	1.0 (2.0)	.86
Total number of face-to-face sessions (at clinic or VC), mean (SD)	3.8 (2.2)	2.9 (2.4)	.24
Total phone sessions, median (IQR)	0.0 (1.0)	1.0 (2.0)	.18
Sessions with child attending, median (IQR)	0.0 (3.0)	0.0 (2.0)	.71

[a]Continuous variables presented as mean (SD) are compared with 2-tailed unpaired *t* tests; variables presented as median (IQR) are compared with Mann-Whitney *U* tests.

[b]TAU-VC: treatment as usual plus videoconferencing.

[c]TAU: treatment as usual.

[d]IQR: interquartile range.

Table 3. Estimated marginal means and treatment effect size for all outcomes based on linear mixed models, adjusted for baseline scores.

Outcome	Estimated marginal 3-month total score, mean (SD)		*F* test (*df*)	Treatment effect size (95% CI)
	TAU-VC[a]	TAU[b]		
Edinburgh Postnatal Depression Scale	11.7 (4.5)	12.1 (4.5)	.053 (22)	–0.42 (–4.23 to 3.91)
Generalized Anxiety Disorder 7-item	8.7 (4.5)	9.1 (4.6)	.050 (21)	–0.44 (–4.49 to 3.62)
Parental Stress Scale	47.6 (6.2)	44.2 (6.2)	1.89 (22)	3.42 (–1.74 to 8.59)

[a]TAU-VC: treatment as usual plus videoconferencing.

[b]TAU: treatment as usual.

Similarly, there were no significant differences in 3-month EPDS, GAD-7, or PSS between groups (see Table 3). When analyses were restricted to those with baseline EPDS >12 (n=10 in the TAU-VC group and n=9 in the TAU group), there were again no significant differences between groups for 3-month EPDS ($F_{1,9}$=0.058, *P*=.81), GAD-7 ($F_{1,9}$=0.004, *P*=.95), or PSS ($F_{1,9}$=0.003, *P*=.96).

Discussion

Principal Findings

In this pilot study, we demonstrated feasibility of recruitment and retention of participants from a program providing psychotherapy to postpartum women with mood and/or anxiety symptoms into a study offering VC as an adjunct to office-based care. We also showed that participants used the intervention, and they reported high satisfaction and substantial savings in time and money. We found no statistically significant differences in total number of sessions attended, and although the number of participants was small, there was no evidence that symptom outcomes differed between those with the option of VC compared with those attending office-based appointments only. Uptake of VC in the intervention group was quite variable. In this study, we offered VC as an option to attending in-person sessions at the clinic, and 74% of the participants used it at least once. Participants used VC for 50% of the sessions on average, with some participants using it often and others choosing not to use it at all. Variability in the uptake of virtual care tools is commonly observed, and it may be attributed to differences in patient-perceived barriers or preferences [27,28]. Among those who used VC, overall, participants were satisfied with the technology. High scores were provided across all the domains of the TSQ, including quality of care provided, similarity to in-person face-to-face interaction, and perception of the interaction. This finding is consistent with other studies that demonstrate high levels of patient acceptance and satisfaction with virtual mental health care [29,30], even when they experience frustration because of technical difficulties [16]. Participants' satisfaction may be associated with their ability to choose the format to receive psychotherapy week to week, increasing their perception of autonomy [31]. Participants subjectively felt that access to optional VC allowed them to attend therapy sessions more frequently. We would hypothesize that by overcoming barriers to care, VC would increase attendance. However, this finding was not statistically observed between groups. A larger study would be better able to determine if access to optional VC affects overall therapy attendance or if access to optional VC affects therapy attendance for certain subgroups.

There appeared to be other advantages of VC as perceived by our participants. Those who used VC reported an average cost savings of Can $26 and time savings of 2.5 hours per session. Over numerous sessions, characteristic of a course of psychotherapy, this amount would add up to significant savings in time and money. To our knowledge, this is the first study to explore patient-rated costs for the use of VC in psychotherapy. Although participants were satisfied with VC, only 1 participant indicated a willingness to pay for VC access out of pocket if it

XSL•FO
RenderX

were not funded by the hospital. This may not be surprising, given that our program exists in a government-funded health care system where patients do not normally have to pay out of pocket for any ambulatory care programs situated in a hospital. Although the sample size was small, it was reassuring that participants who used VC had similar outcomes to those who received TAU. Sometimes, when technology is leveraged, the increased convenience and ease of access have to be weighed against a possible reduction in treatment efficacy [32]. Although a larger study would be required to generate more certainty about the outcomes, it is promising that it does not appear to be the case for a mental health treatment model, where in-person care is supplemented by virtual care for postpartum women.

Limitations

This pilot study used a sample of 38 participants recruited from an academic ambulatory hospital in an urban city; therefore, its patients may have access to more resources than other parts of the province where access barriers are more pronounced. Our program is unique in that it offers specialized care in an ambulatory hospital that provides access to free childcare during appointments, which eliminates some of the barriers to care in the postpartum population. In addition, we have access to a secure government-funded VC platform. In this setting, we evaluated the option of VC for treatment, which in some areas may not be feasible when there is no access to office-based treatment and VC is the only option. Over 80% of the participants in the study had had some therapy experience before being referred to our program; therefore, the results may not be as generalizable to a therapy-naïve cohort. That said, we demonstrated the feasibility of recruiting and retaining postpartum women receiving psychotherapy for mood and anxiety difficulties to receive the option of VC in addition to office-based care. In addition, acceptability of the intervention was high, and there were no indications of significant differences in outcomes, supporting that a larger study is warranted.

Conclusions

This novel research protocol explored the acceptability and preliminary effectiveness of VC in combination with a treatment-as-usual protocol of in-person and telephone-based care for psychotherapy during the postpartum period within an urban center. VC was perceived as an acceptable format in which to receive psychotherapy, with patient-reported cost and time savings. Virtually available options to help patients overcome barriers to care can enhance the patient-centeredness of ambulatory care, most likely without any compromise to patient outcomes in this population. The results herein support proceeding to a larger study to definitively evaluate this protocol for its efficacy in this vulnerable population. Future work should also evaluate the use of VC in other urban settings and patient populations.

Acknowledgments

The authors would like to thank the postpartum women who participated in the study and the psychotherapists who provided care to these women. This study was funded by a project grant awarded to JMH from the Women's College Hospital Academic and Medical Services Group/Association of Fundraising Professionals Innovation Fund.

Conflicts of Interest

SNV receives royalties from UpToDate for chapters authored on depression and pregnancy. The authors have no other competing interests to declare.

Editorial note: This randomized study was not registered, justified by the authors as it was a pilot study. The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to their primary outcomes or effectiveness, as the lack of registration means that authors could change their outcome measures retrospectively.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 561KB - jmir_v21i6e13172_app1.pdf]

References

1. Mojtabai R, Olfson M, Sampson N, Jin R, Druss B, Wang P, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. Psychol Med 2011 Aug;41(8):1751-1761. [doi: 10.1017/S0033291710002291] [Medline: 21134315]
2. Vasiliadis H, Lesage A, Adair C, Boyer R. Service use for mental health reasons: cross-provincial differences in rates, determinants, and equity of access. Can J Psychiatry 2005 Sep;50(10):614-619. [doi: 10.1177/070674370505001007] [Medline: 16276852]
3. Barr J. Postpartum depression, delayed maternal adaptation, and mechanical infant caring: a phenomenological hermeneutic study. Int J Nurs Stud 2008 Mar;45(3):362-369. [doi: 10.1016/j.ijnurstu.2006.10.002] [Medline: 17157848]

XSL•FO
RenderX

4. Pearson R, Cooper R, Penton-Voak I, Lightman S, Evans J. Depressive symptoms in early pregnancy disrupt attentional processing of infant emotion. Psychol Med 2010 Apr;40(4):621-631. [doi: 10.1017/S0033291709990961] [Medline: 19671214]
5. Goodman J. Women's attitudes, preferences, and perceived barriers to treatment for perinatal depression. Birth 2009 Mar;36(1):60-69. [doi: 10.1111/j.1523-536X.2008.00296.x] [Medline: 19278385]
6. Maloni J, Przeworski A, Damato E. Web recruitment and internet use and preferences reported by women with postpartum depression after pregnancy complications. Arch Psychiatr Nurs 2013 Apr;27(2):90-95. [doi: 10.1016/j.apnu.2012.12.001] [Medline: 23540519]
7. Haga S, Ulleberg P, Slinning K, Kraft P, Steen T, Staff A. A longitudinal study of postpartum depressive symptoms: multilevel growth curve analyses of emotion regulation strategies, breastfeeding self-efficacy, and social support. Arch Womens Ment Health 2012 Jun;15(3):175-184. [doi: 10.1007/s00737-012-0274-2] [Medline: 22451329]
8. Meyer B, Berger T, Caspar F, Beevers C, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009 May 11;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
9. Foulkes M. Enablers and barriers to seeking help for a postpartum mood disorder. J Obstet Gynecol Neonatal Nurs 2011;40(4):450-457. [doi: 10.1111/j.1552-6909.2011.01264.x] [Medline: 21771070]
10. Nielsen Forman D, Videbech P, Hedegaard M, Dalby Salvig J, Secher N. Postpartum depression: identification of women at risk. BJOG 2000 Oct;107(10):1210-1217 [FREE Full text] [Medline: 11028570]
11. Jamieson T, Wallace R, Armstrong K, Agarwal P, Griffin B, Wong I, et al. Women's College Hospital. 2015. Virtual care: a framework for a patient-centric system URL: https://www.womenscollegehospital.ca/assets/pdf/wihv/WIHV_VirtualHealthSymposium.pdf [accessed 2019-05-09] [WebCite Cache ID 78FS22u7q]
12. Osenbach J, O'Brien KM, Mishkind M, Smolenski D. Synchronous telehealth technologies in psychotherapy for depression: a meta-analysis. Depress Anxiety 2013 Nov;30(11):1058-1067. [doi: 10.1002/da.22165] [Medline: 23922191]
13. Tarp K, Bojesen A, Mejldal A, Nielsen A. Effectiveness of optional videoconferencing-based treatment of alcohol use disorders: randomized controlled trial. JMIR Ment Health 2017 Sep 29;4(3):e38 [FREE Full text] [doi: 10.2196/mental.6713] [Medline: 28963093]
14. Erbe D, Eichert H, Riper H, Ebert D. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Dec 15;19(9):e306 [FREE Full text] [doi: 10.2196/jmir.6588] [Medline: 28916506]
15. Thorp S, Fidler J, Moreno L, Floto E, Agha Z. Lessons learned from studies of psychotherapy for posttraumatic stress disorder via video teleconferencing. Psychol Serv 2012 May;9(2):197-199. [doi: 10.1037/a0027057] [Medline: 22662733]
16. Backhaus A, Agha Z, Maglione M, Repp A, Ross B, Zuest D, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv 2012 May;9(2):111-131. [doi: 10.1037/a0027924] [Medline: 22662727]
17. Power A, Gordon A. There's an app for that—but how do we know if it's a good one? Br J Midwifery 2015;9(2):442-444. [doi: 10.12968/bjom.2015.23.6.442]
18. Sawyer M, Reece C, Bowering K, Jeffs D, Sawyer A, Mittinty M, et al. Nurse-moderated internet-based support for new mothers: non-inferiority, randomized controlled trial. J Med Internet Res 2017 Dec 24;19(7):e258 [FREE Full text] [doi: 10.2196/jmir.6839] [Medline: 28739559]
19. Rai M, Vigod S, Hensel J. Barriers to office-based mental health care and interest in e-communication with providers: a survey study. JMIR Ment Health 2016 Aug 01;3(3):e35 [FREE Full text] [doi: 10.2196/mental.6068] [Medline: 27480108]
20. Rees C, Maclaine E. A systematic review of videoconference-delivered psychological treatment for anxiety disorders: videoconferenced treatment for anxiety disorders. Aust Psychol 2015;50(4):259-264. [doi: 10.1111/ap.12122]
21. Hertzog M. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
22. Gibson J, McKenzie-McHarg K, Shakespeare J, Price J, Gray R. A systematic review of studies validating the Edinburgh Postnatal Depression Scale in antepartum and postpartum women. Acta Psychiatr Scand 2009 May;119(5):350-364. [doi: 10.1111/j.1600-0447.2009.01363.x] [Medline: 19298573]
23. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
24. Berry J, Jones W. The parental stress scale: initial psychometric evidence. J Soc Pers Relatsh 1995;12(3):263-272. [doi: 10.1177/0265407595123009]
25. Shapiro D, Stewart A. Parenting stress, perceived child regard, and depressive symptoms among stepmothers and biological mothers. Fam Relat 2011;60(5):533-544. [doi: 10.1111/j.1741-3729.2011.00665.x]
26. Yip M, Chang A, Chan J, MacKenzie A. Development of the Telemedicine Satisfaction Questionnaire to evaluate patient satisfaction with telemedicine: a preliminary study. J Telemed Telecare 2003;9(1):46-50. [doi: 10.1258/135763303321159693] [Medline: 12641893]

XSL•FO
RenderX

27. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Dec 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
28. Hollis C, Falconer C, Martin J, Whittington C, Stockton S, Glazebrook C, et al. Annual research review: digital health interventions for children and young people with mental health problems - a systematic and meta-review. J Child Psychol Psychiatry 2017 Apr;58(4):474-503. [doi: 10.1111/jcpp.12663] [Medline: 27943285]
29. Richardson L, Frueh BC, Grubaugh A, Egede L, Elhai J. Current directions in videoconferencing tele-mental health research. Clin Psychol (New York) 2009 Sep 01;16(3):323-338 [FREE Full text] [doi: 10.1111/j.1468-2850.2009.01170.x] [Medline: 20161010]
30. Grady B, Myers K, Nelson E, Belz N, Bennett L, Carnahan L, American Telemedicine Association Telemental Health Standards and Guidelines Working Group. Evidence-based practice for telemental health. Telemed J E Health 2011 Mar;17(2):131-148. [doi: 10.1089/tmj.2010.0158] [Medline: 21385026]
31. Tarp K, Nielsen A. Patient perspectives on videoconferencing-based treatment for alcohol use disorders. Alcohol Treat Q 2017;35(4):344-358. [doi: 10.1080/07347324.2017.1348785]
32. Martinez K, Rood M, Jhangiani N, Kou L, Boissy A, Rothberg M. Association between antibiotic prescribing for respiratory tract infections and patient satisfaction in direct-to-consumer telemedicine. JAMA Intern Med 2018 Nov 01;178(11):1558-1560. [doi: 10.1001/jamainternmed.2018.4318] [Medline: 30285050]

Abbreviations

EPDS: Edinburgh Postnatal Depression Scale
GAD-7: Generalized Anxiety Disorder 7-item
PRCQ: Patient Reported Costs Questionnaire
PSS: Parental Stress Scale
TAU: treatment as usual
TAU-VC: treatment as usual plus videoconferencing
TSQ: Telemedicine Satisfaction Questionnaire
VC: videoconferencing

Edited by C Eichenberg; submitted 18.12.18; peer-reviewed by S Brigitte, W Bennett; comments to author 14.01.19; revised version received 11.03.19; accepted 07.04.19; published 20.05.19.

Please cite as:
Yang R, Vigod SN, Hensel JM
Optional Web-Based Videoconferencing Added to Office-Based Care for Women Receiving Psychotherapy During the Postpartum Period: Pilot Randomized Controlled Trial
J Med Internet Res 2019;21(6):e13172
URL: https://www.jmir.org/2019/6/e13172/
doi:10.2196/13172
PMID:31199291

©Rebecca Yang, Simone N Vigod, Jennifer M Hensel. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Tracking Healthy People 2020 Internet, Broadband, and Mobile Device Access Goals: An Update Using Data From the Health Information National Trends Survey

Alexandra J Greenberg-Worisek[1], PhD, MPH; Shaheen Kurani[1], MSc; Lila J Finney Rutten[1], PhD, MPH; Kelly D Blake[2], SCD; Richard P Moser[2], PhD; Bradford W Hesse[2], PhD

[1]Mayo Clinic College of Medicine and Science, Rochester, MN, United States

[2]National Cancer Institute, National Institutes of Health, Rockville, MD, United States

Corresponding Author:
Alexandra J Greenberg-Worisek, PhD, MPH
Mayo Clinic College of Medicine and Science
200 First Street SW
Rochester, MN, 55906
United States
Phone: 1 507 538 7388
Email: greenberg.alexandra@mayo.edu

Abstract

Background: As the year 2020 approaches, there is a need to evaluate progress toward the United States government's Healthy People 2020 (HP2020) health information technology and communication objectives to establish baselines upon which Healthy People 2030 objectives can be based.

Objective: The aim of this study was to use the National Cancer Institute's (NCI) Health Information National Trends Survey (HINTS) to benchmark progress toward HP2020 goals related to increasing internet access using broadband, and to assess the state of the digital divide for various sociodemographic groups.

Methods: We merged and analyzed data from 8 administrations of HINTS (2003-2017). Descriptive statistics were generated, and predicted marginals were calculated using interaction terms between survey year and selected sociodemographic variables of interest, including age, sex, race and ethnicity, income, education, and geography (rural versus urban), to test for differential change over time.

Results: The number of users having access to the internet increased between 2003 and 2014 (63.15% [3982/6358] to 83.41% [2802/3629]); it remained relatively steady from 2014 to 2017 (81.15% [2533/3283]). Broadband access increased between 2003 and 2011 (from 32.83% [1031/3352] to 77.87% [3375/4405]), but has been declining since (55.93% [1364/2487] in 2017). Access via cellular network increased between 2008 and 2017 (from 6.86% [240/4405] to 65.43% [1436/2489]). Statistically significant disparities in overall internet access were noted in the predicted marginals for age, sex, race and ethnicity, income, and education; for age, sex, income, and geography for broadband access; and for age and sex for cellular network.

Conclusions: The targets set forth in HP2020 were met for overall internet access and for internet access via cellular network; however, the target was not met for internet access via broadband. Furthermore, although the digital divide persisted by sociodemographic characteristics, the magnitude of many disparities in access decreased over time.

(J Med Internet Res 2019;21(6):e13300) doi:10.2196/13300

KEYWORDS

Healthy People 2020; digital divide; internet

Introduction

The Healthy People initiative sets 10-year objectives for improving the health of Americans nationwide based on the latest scientific evidence [1]. Some Healthy People objectives focus on communication-related objectives including access to information and communication resources such as the internet and mobile devices. These objectives related to technology access speak to their increasing importance in managing health and health care in the United States.

XSL•FO
RenderX

The Office of Disease Prevention and Health Promotion (ODPHP) selects relevant and scientifically rigorous data sources to benchmark progress toward objectives outlined in Healthy People programs. For the Health Communication and Health Information Technology (HC/HIT) objectives related to internet access, ODPHP chose items regularly included in the National Cancer Institute (NCI)'s Health Information National Trends Survey (HINTS), a nationally representative, probability-based survey whose primary aims are to track health behaviors, communication, and technology use [2].

In Healthy People 2020 (HP2020), one of the HC/HIT objectives is to *increase the proportion of individuals with access to the internet* to 75.4%, a 10% improvement over the percentage observed in 2007 (68.5%) [3]. There are 2 additional subobjectives within this broader objective of increasing overall internet access: (1) to increase internet access via broadband by 10% (from 75.6% of those with internet access in 2007 [HP2020 Target: 83.2%]) and (2) to increase internet access via mobile by 10% (from 6.7% of those with internet access in 2007 [HP2020 Target: 7.4%]).

As the Healthy People initiative enters its fourth decade and fourth iteration with the upcoming Healthy People 2030 (HP2030) objectives, it is necessary to reflect upon the HP2020 objectives and assess whether these targets were met and whether these objectives remain relevant in the context of a rapidly evolving communication technology landscape [4]. We analyzed recent HINTS data (2017) to update our previous report on progress toward these HC/HIT objectives and to examine current disparities in access to the internet via broadband and cellular network [5]. In addition, we examine the impact of geography (urban vs rural residence) on internet access via these different connections.

Methods

Survey Population and Data Collection

Data from 8 administrations of the NCI's HINTS were merged for these analyses (N=37,415; Table 1; expanded from Serrano et al [5]). Briefly, HINTS is a national cross-sectional survey of US adults that collects data from the public on a broad range of health and cancer information, communication, attitudes and behaviors, and use of health information technologies. HINTS uses a probability-based sampling frame with a 2-level design in which residential addresses in the United States are sampled, and then 1 adult from each address is randomly selected for participation. In later administrations of HINTS (HINTS 4 and later), efforts were made specifically to oversample for those residing in central Appalachia as well as minority populations. For additional details, please see the corresponding methodology reports for each administration [2,6-8].

Table 1. Details of the 8 survey administrations of Health Information National Trends Survey, during 2003-2018 (N=37,415).

Variable	HINTS[a] 1 (2003)	HINTS 2 (2005)	HINTS 3 (2007-2008)	HINTS 4 Cycle 1 (2011)	HINTS 4 Cycle 2 (2012)	HINTS 4 Cycle 3 (2013)	HINTS 4 Cycle 4 (2014)	HINTS 5 Cycle 1 (2017)
Survey period	Oct 2002-Apr 2003	Feb 2005-Aug 2005	Jan 2008-May 2008	Oct 2011-Jan 2012	Oct 2012-Dec 2012	Sept 2013-Dec 2013	Aug 2014-Nov 2014	Jan 2017-Mar 2017
Respondents (n)	6369	5586	Mail: 3582; RDD[b]: 4092	3959	3630	3185	3677	3335
Survey mode	RDD	RDD	Mail and RDD	Mail	Mail	Mail	Mail	Mail
Response rate (%)	33.1	20.8	Mail: 40.0; RDD: 24.2	36.7	40.0	35.2	34.4	32.4

[a]HINTS: Health Information National Trends Survey.

[b]RDD: random digit dialing.

Dependent Variables

Our primary outcome variable of interest was internet access; the specific survey item used was *Do you ever go on-line to access the Internet or the World Wide Web, or to send and receive e-mail?* Wording of this item was consistent across survey administrations.

The second survey item of interest was aimed at assessing progress on the HP2020 goals for broadband access. This item was asked in each of the 8 survey administrations; however, wording of this item changed throughout the survey administrations as technology advanced. The wording for HINTS 1-3 was as follows: "When you use the internet at home, do you mainly access it through…cable or satellite modem?" and "…a DSL modem?" These 2 items were combined into 1 variable, to better align with the wording in HINTS 4 [all cycles] and HINTS 5 Cycle 1: "…broadband such as DSL [digital subscriber line], cable, or FiOS [fiber optic service]?"

Finally, the third HP2020 goal examined was for cellular internet access. To examine cellular internet access, the following item was examined: "When you use the Internet do you access it through…a cellular network?" [Yes/No; HINTS 3, HINTS 4 Cycles 1-4, HINTS 5 Cycle 1].

Independent Variables

Included in the analyses were sociodemographic variables shown to be related to the digital divide (age, sex, race and ethnicity, education, income, and geography) [9-11]. Age was analyzed categorically rather than continuously, as was income. Sex responses were either male or female. Education was categorized as less than high school, high school, some college, and college graduate or higher. Geography was dichotomized into rural and urban using the US Department of Agriculture's Rural-Urban

XSL•FO
RenderX

Continuum Codes (RUCCs). Urban categorization included RUCCs 1-3, which represent metro area counties with greater than 20,000 residents; rural categorization included RUCCs 4-9, which represent nonmetro counties with populations ranging from 2500 to 20,000 [12].

Statistical Analyses

Analyses were conducted using SAS 9.4 survey procedures to accommodate the complex sampling procedure used and incorporate the jackknife replicate weights. All analyses were weighted to produce population-level point estimates. Descriptive statistics were calculated for all items. In addition, sociodemographic factors were analyzed using logistic regression analyses with predicted marginals to determine whether there existed significant differences in groups for each of the outcome variables. Interaction terms were included between each independent variable and survey year to investigate differential change over survey years.

Results

Overview

Although respondent characteristics varied slightly across survey administrations, the weighted respondent characteristics closely reflected those of the US Census for each respective iteration. As expected in most mailed survey studies, most respondents were female, aged 18 to 34, non-Hispanic white, had at least some college education, and had a household income of US $75,000 or more annually (data not shown; available on the website [13]).

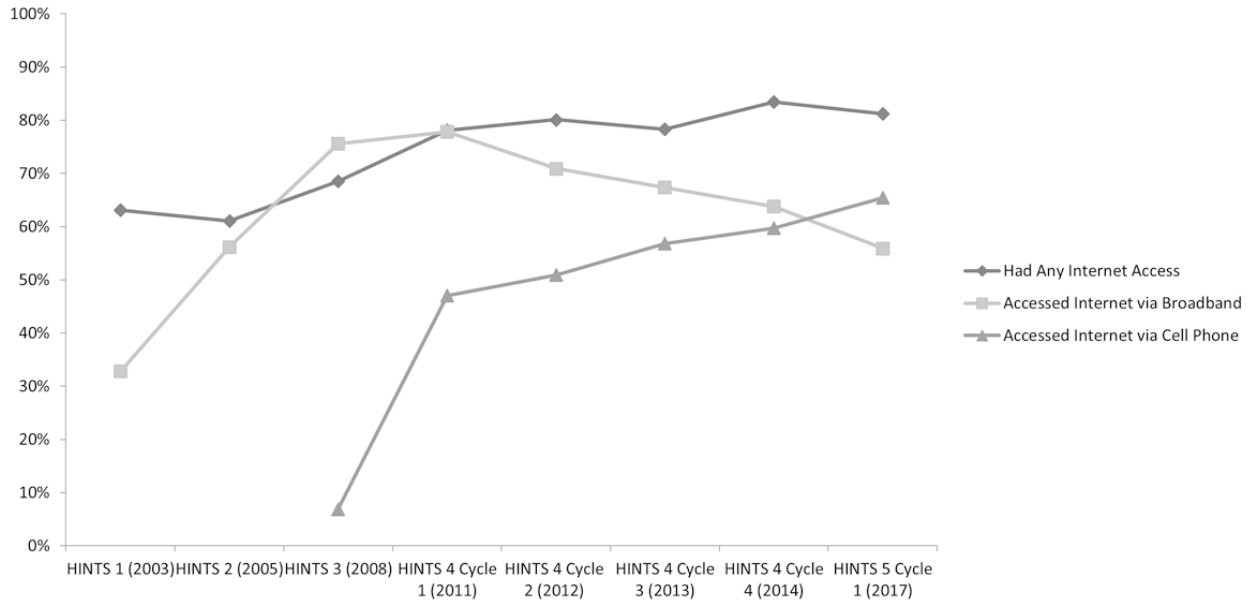
Internet Access

Internet access increased overall between 2003 and 2014 (20.3 percentage points, from 63.15% [3982/6358] to 83.41% [2802/3629]), with some variation along the way; however, the percentage of those with internet access remained relatively steady from 2014 to 2017 (Figure 1).

In the multivariable regression model, all of the sociodemographic variables examined showed statistical significance between groups after adjusting for survey year; however, the magnitude of differences was not significant across all groups within each variable (Multimedia Appendix 1). Briefly, men had a 0.74-fold decreased odds of having internet access as compared with women (95% CI 0.67-0.83). All age categories 35 years and above had significantly lower odds of having internet access compared with those in the 18 to 34 years old referent group; these odds ranged from 0.42 in the 35 to 49 year old age group (95% CI 0.35-0.51) to 0.04 in the 75+ age group (95% CI 0.03-0.05). Every race and ethnicity group was significantly less likely to have internet access compared with the non-Hispanic white referent group; Hispanic respondents had a 0.37 odds (95% CI 0.31-0.43), non-Hispanic Black had 0.51 odds (95% CI 0.44-0.59), and non-Hispanic Other had 0.45 (95% CI 0.34-0.61).

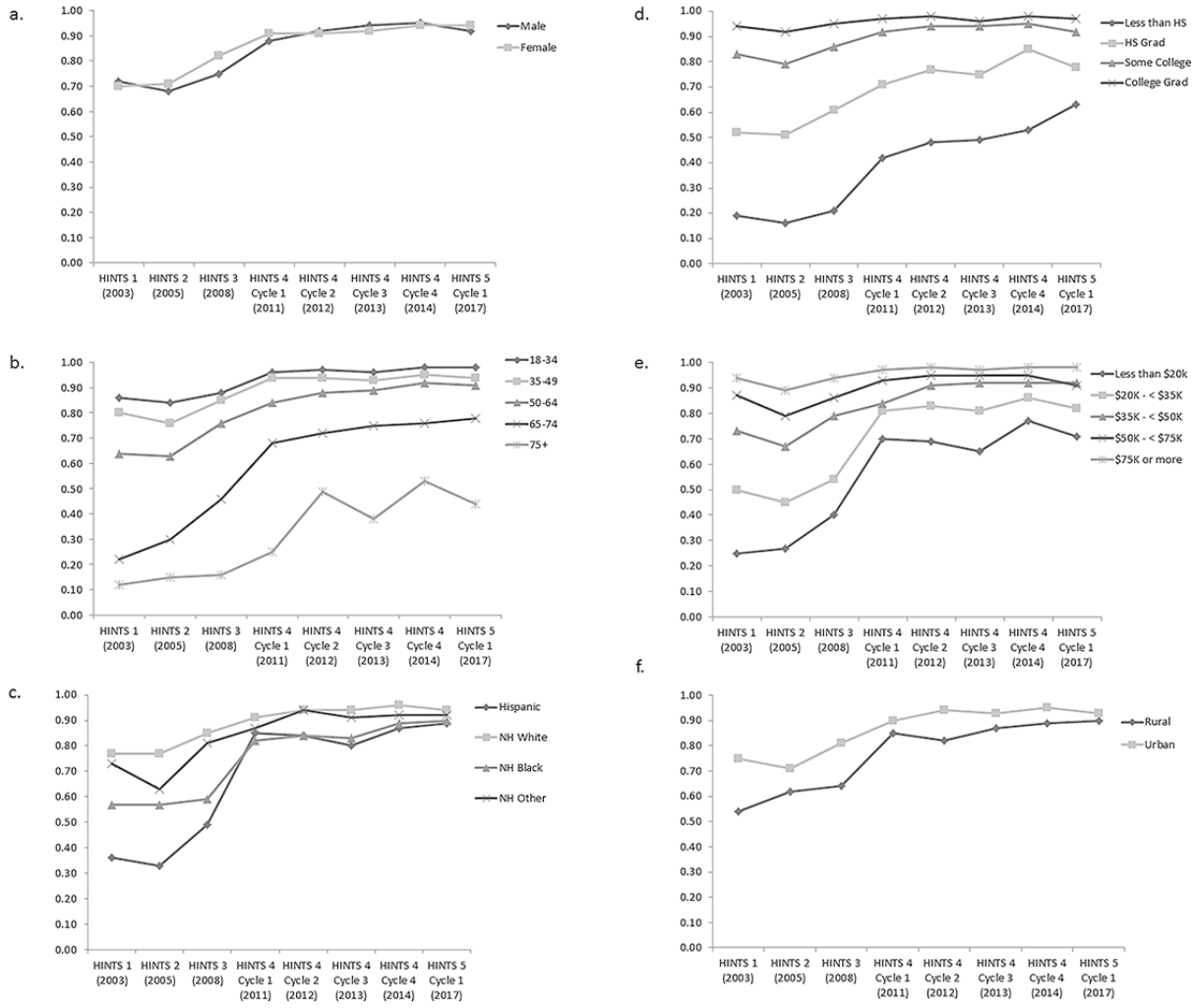
Those with higher levels of education had significantly higher odds of having internet access versus those with less than a high school education; for example, college graduates had an increased odds of having internet access compared with those who did not graduate from high school (odds ratio [OR] 9.44; 95% CI 7.65-11.65). Similarly, those with higher annual incomes had increased odds compared with those with lower annual incomes (ie, those with US $75,000 or higher annual income had 6.67-fold increased odds of having internet access compared with those with less than US $20,000 per year income [95% CI 5.54, 8.04]). Finally, those living in rural areas had reduced odds of having internet access compared with those residing in urban areas (OR 0.75; 95% CI 0.67-0.84).

Figure 1. Percentage of US adult population with access to internet (out of all respondents), access to internet via broadband (out of respondents with internet access), and access to internet via cell phone (out of respondents with internet access), Health Information National Trends Survey (HINTS) 2003-2017. Survey question on accessing the internet through a cellular network not included in HINTS 1 (2003) or HINTS 2 (2005) survey administrations.



XSL•FO
RenderX

Figure 2. Trends of having internet access based on responses from the National Cancer Institute's Health Information National Trends Survey administrations between 2003 and 2017. (a) Predicted marginals by sex. (b) Predicted marginals by age. (c) Predicted marginals by race and ethnicity. (d) Predicted marginals by education. (e) Predicted marginals by income. (f) Predicted marginal by geography. All models adjusted for sex, age, race and ethnicity, education, income, and geography. NH: non-Hispanic.



Although all sociodemographic variables examined were statistically significant in our multivariable logistic regression, age category, sex, education, income, and race and ethnicity had significant interactions with survey year (Figure 2). The interaction between survey year and geography had borderline statistical significance (P=.07). For most groups examined within each of these variables, the overall trend was toward increasing internet access over time.

Internet Access via Broadband

Broadband access increased overall by 42.8 percentage points from 2003 to 2011 (from 32.83% [1031/3352] to 77.87% [3375/4405]); however, it has continually decreased since then (to 55.93% [1364/2487] in 2017, 21.9 percentage points; Figure 1). It should be noted that broadband access from the perspective of the HINTS survey emphasizes wired access to the home (eg, through FiOS, cable, and DSL). It will be distinguished for the purposes of this paper from cellular access to the internet as provided through mobile phone technologies and mobile data plans.

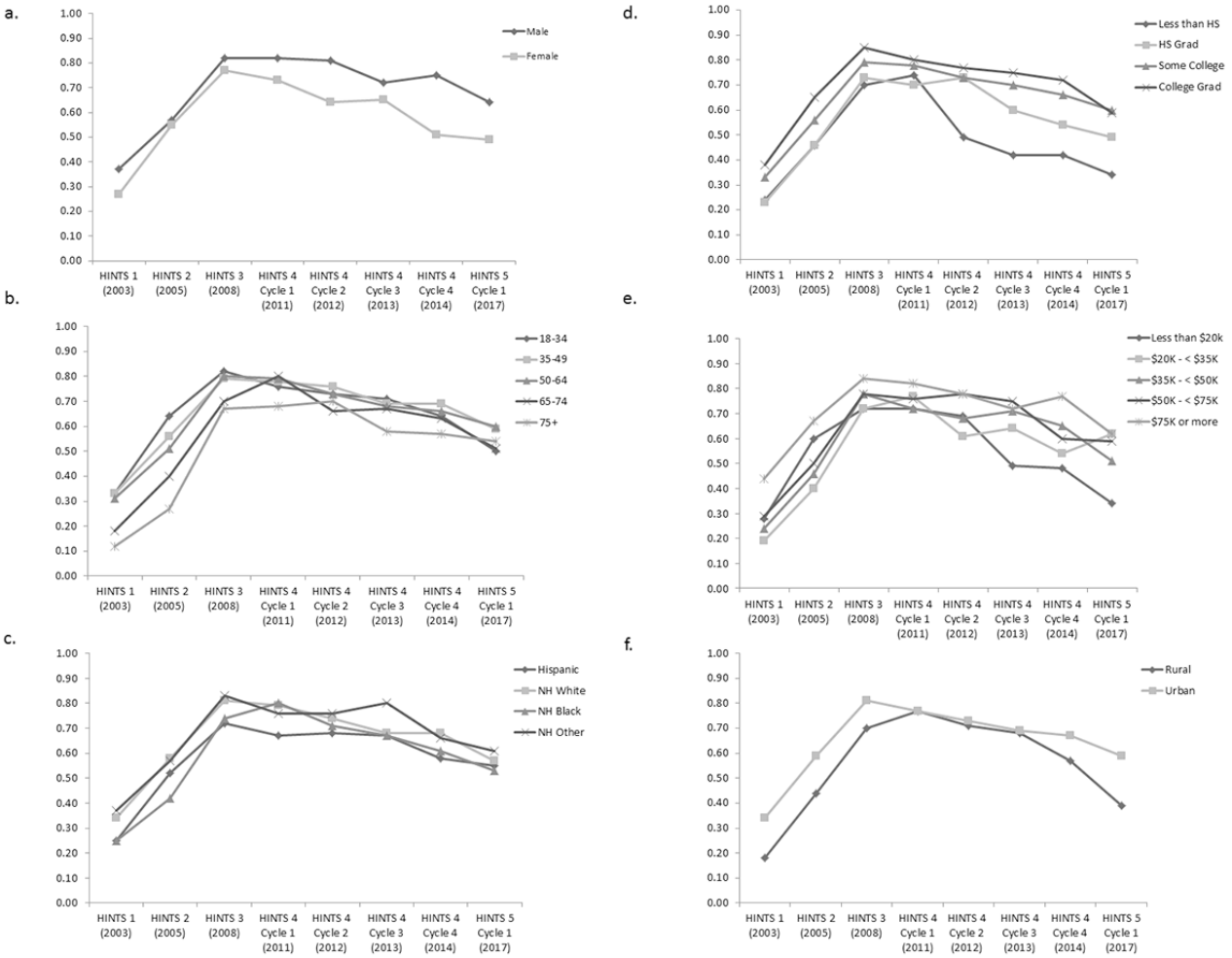
In the multivariable logistic regression model of having broadband access among those with internet access, all sociodemographic variables were again statistically significant after adjusting for survey year; however, the magnitude of some of these differences was greatly reduced (Multimedia Appendix 2). Although most trends within variables remained the same, there were a few notable changes. First, males who reported having internet access had a 1.67-fold increased odds of having broadband access as compared with their female counterparts (95% CI 1.51-1.85). The only significant difference among race and ethnicity groups was the reduced odds of having broadband among Hispanic respondents reporting having internet access as compared with non-Hispanic white respondents (OR 0.77; 95% CI 0.63-0.94); there was no significant difference for non-Hispanic blacks and non-Hispanic other compared with non-Hispanic whites, suggesting they are no more or less likely to have broadband internet access. Within education categories, there was no significant difference between high school graduates as compared with nonhigh school graduates (OR 1.33; 95% CI 0.93-1.91); differences remained significant for those who had some college or who were college graduates, compared

XSL•FO
RenderX

with those with less than a high school diploma (OR 1.84; 95% CI 1.32-2.55 for those with some college and OR 1.97; 95% CI 1.42-2.74 for those with college graduates). Similarly, differences were no longer statistically significant between those with annual incomes of US $20,000-50,000 compared with those of less than US $20,000 per year; they remained significantly different for higher income levels (OR 1.45, 95% CI 1.14-1.84 for US $50,000 to < US $75,000; OR 1.85, 95% CI 1.55-2.21 for > US $75,000; Multimedia Appendix 2).

Although all sociodemographic variables examined were statistically significant in our multivariable logistic regression, only age category, sex, income, and geography showed statistically significant interactions with survey year (Figure 3), suggesting that access changed over time for these groups. For most groups examined within each of the sociodemographic variables, the overall trend was toward increasing internet access via broadband until 2011, with a subsequent decrease after that year.

Figure 3. Trends of having internet access via broadband based on responses from the National Cancer Institute's Health Information National Trends Survey administrations between 2003 and 2017. (a) Predicted marginals by sex. (b) Predicted marginals by age. (c) Predicted marginals by race and ethnicity. (d) Predicted marginals by education. (e) Predicted marginals by income. (f) Predicted marginal by geography. All models adjusted for sex, age, race and ethnicity, education, income, and geography. NH: non-Hispanic.



Internet Access via Cellular Network

Internet access via cellular network increased from 2008 to 2017 (58.5 percentage points, from 6.86% [240/4405] to 65.43% [1436/2489]); the greatest increase was between 2008 and 2011 (40.1 percentage points, from 6.86% [240/4405] to 47.01% [1128/2861]; Figure 1). This finding presents a contrasting trend to the recent decline in internet access through traditional landline, fiber optic, or cable broadband to the home.

Most of the sociodemographic variables within our multivariable model were statistically significant after adjusting for survey year, save for geography (Table 2). No significant difference was found between males and females in accessing the internet via cellular networks, nor for rural versus urban residents. There was no significant difference in having internet access via cellular network for Hispanics and non-Hispanic blacks compared with non-Hispanic whites. Similarly, there was no statistically significant difference between educational groups of high school graduates and above compared with those who had less than a high school education. Odds of accessing the internet via cellular network decreased with increasing age in a statistically significant manner for each age category above 35 years of age, compared with those 18 to 34 years of age.

XSL•FO
RenderX

Differences were not significant between those with annual incomes of US $20,000 to 50,000 compared with those of less than US $20,000 per year, but again remained significantly different for higher income levels (Table 2).

Table 2. Weighted multivariate logistic regression model of predictors of having internet access via mobile phone among those who reported having internet access. Data from the National Cancer Institute's Health Information National Trends Survey (HINTS) administrations between 2008 and 2017 (n=14,794).

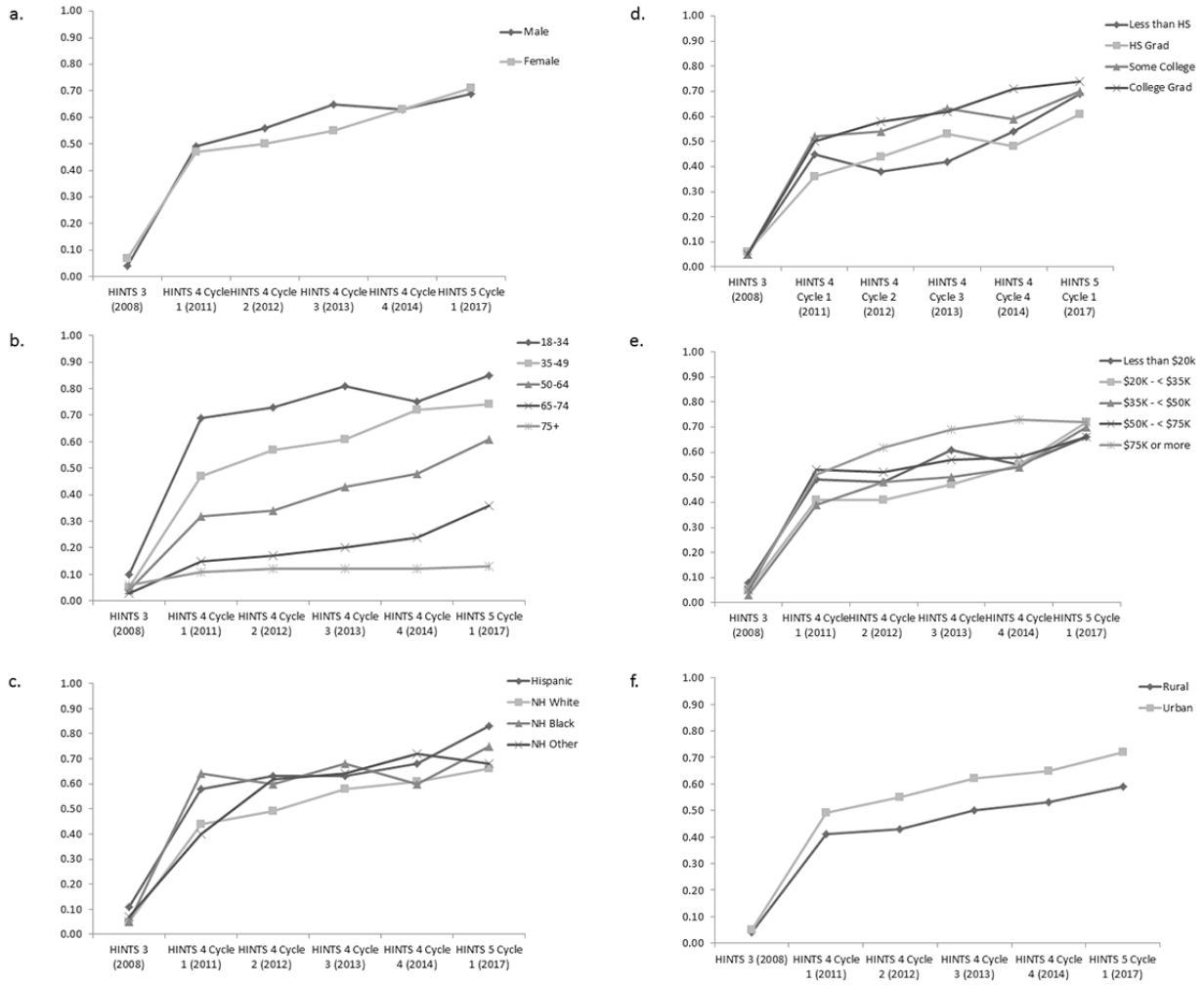
Variable	Predictors of internet access via cell phone				
	Odds ratio (95% CI)	Beta coefficient	SE beta	Adjusted Wald F	*P* value
Sex				1.76	<.001
Female	Ref[a]	Ref	Ref		
Male	1.09 (0.96-1.24)	0.08	0.06		
Age				951.63	<.001
18-34	Ref	Ref	Ref		
35-49	0.45 (0.38-0.53)	–0.81	0.08		
50-64	0.21 (0.18-0.25)	–1.57	0.08		
65-74	0.08 (0.07-0.10)	–2.49	0.11		
>75	0.04 (0.03-0.06)	–3.19	0.15		
Race and ethnicity				9.90	.02
Non-Hispanic White	Ref	Ref	Ref		
Hispanic	1.28 (1.02-1.61)	0.25	0.11		
Non-Hispanic Black	1.36 (1.06-1.75)	0.31	0.12		
Non-Hispanic Other	0.84 (0.62-1.12)	–0.17	0.15		
Education				10.91	.012
Less than high school	Ref	Ref	Ref		
High school graduate	1.03 (0.61-1.75)	0.03	0.26		
Some college	1.41 (0.81-2.46)	0.35	0.28		
College graduate	1.43 (0.84-2.42)	0.36	0.26		
Income (US $)				102.94	<.001
<$20,000	Ref	Ref	Ref		
$20,000 to <$35,000	0.98 (0.70-1.37)	–0.02	0.17		
$35,000 to <$50,000	1.07 (0.82-1.41)	0.07	0.13		
$50,000 to <$75,000	1.34 (1.06-1.71)	0.30	0.12		
$75,000 +	1.93 (1.56-2.38)	0.66	0.11		
Geography				3.17	.08
Urban	Ref	Ref	Ref		
Rural	0.82 (0.66-1.03)	–0.19	0.11		
HINTS[b] Survey Year				648.91	<.001
HINTS 3 (2008)	Ref	Ref	Ref		
HINTS 4 Cycle 1 (2011)	17.71 (13.16-23.84)	2.87	0.15		
HINTS 4 Cycle 2 (2012)	21.39 (15.22-30.07)	3.06	0.17		
HINTS 4 Cycle 3 (2013)	29.01 (21.17-39.75)	3.37	0.16		
HINTS 4 Cycle 4 (2014)	31.73 (22.02-45.72)	3.46	0.18		
HINTS 5 Cycle 1 (2017)	50.48 (36.53-69.74)	3.92	0.16		

[a]Ref: reference group.

[b]HINTS: Health Information National Trends Survey.

XSL•FO
RenderX

Figure 4. Trends of having internet access via cell phone/mobile based on responses from the National Cancer Institute's Health Information National Trends Survey administrations between 2008 and 2017. (a) Predicted marginals by sex. (b) Predicted marginals by age. (c) Predicted marginals by race and ethnicity. (d) Predicted marginals by education. (e) Predicted marginals by income. (f) Predicted marginal by geography. All models adjusted for sex, age, race and ethnicity, education, income, and geography. NH: non-Hispanic.



Although most sociodemographic variables examined were statistically significant, only age category and sex had significant interactions with survey year (Figure 4), suggesting that mobile technology may have helped or be helping bridge the digital divide across income, race and ethnicity, education, and geography over time [14,15].

Discussion

Principal Findings

The results presented in this study provide a timely update on progress toward the HP2020 goals related to access to the internet during a time of rapid changes in communication and information technology. The HP2020 objective to increase internet access overall for Americans by 10% from the 2007 baseline percentage of 68.5% (HP2020 Target: 75.4%) [3] was surpassed in 2011 (78.1%), as noted in our 2016 report; it has continued to remain relatively steady since the previous report (81.2% in 2017) [5]. The HP2020 objective to increase internet access via broadband by 10% from the baseline percentage of 75.6% in 2007 (HP2020 Target: 83.2%) was never reached, as it peaked at 77.8% in 2011 and then steadily declined. As stated earlier, we believe that this may be due to the increasing shift toward internet access via cellular network, as the data presented here demonstrate. The HP2020 objective to increase internet access via cellular network by 10% from the baseline of 6.7% in 2007 (HP2020 Target: 7.4%) was greatly surpassed, with 65.4% of individuals in 2017 reporting internet access via cellular network.

Overall, internet access has remained relatively stable over the last 5 years; however, contrasting trends in traditional wired access versus cellular access illustrate an important nuance over the ways in which populations access the internet that was not yet obvious in the previous publications [4,5]. Broadband access to the home offers always-on, high-speed capacity to search for health information, review or download data from a personal health record, to order medications, and so on. The rise of cellular access provides patients with an always-on, always present capacity. It should be noted, then, that the steady state of internet access reflects divergence across channels, with some capabilities common to both (eg, internet-based searching) and with some capabilities ideally suited for one over the other. For

XSL•FO
RenderX

example, some websites that contain detailed information may not be easily viewed on a mobile phone; however, the use of a tablet via cellular network may allow such information to be accessed fully. Of the variables examined, only age has statistically significant interactions over time across internet access overall, internet access via broadband, and internet access via cellular network. This suggests that the digital divide by age persists, wherein older adults are less likely to report overall internet access than those aged 18 to 34.

That disparities in overall internet access and via broadband exist independent of survey administration indicates that a digital divide persists for many; however, our predicted marginals reported here indicate that the magnitude of these divides may be narrowing for some groups [16]. These findings are consistent with studies that show that the introduction of the smartphone in 2007 has helped to narrow gaps in internet access for many over the past ten years [17-19]. That, coupled with the decreasing cost of such technology and data plans, means that internet access via cellular network is within reach for a broader proportion of Americans. Specific groups may benefit from targeted interventions to increase access to and acceptance of the internet, such as members of older age groups who face greater health challenges and may therefore benefit from access to health information and communication technologies [14,20].

Although a significant difference was observed between rural and urban residents in overall internet access and in accessing the internet via broadband among those who reported having internet access, no difference was seen in internet access via cellular network among those reporting having internet access. We hypothesize that this may be explained in part due to the fact that there are fewer infrastructural barriers for expanding internet access via cellular network to rural areas than there are for expanding internet access via broadband, as well as due to the reduction in broadband expansion after 2014 reported by the Federal Communications Commission [21]. Such information can be used to help public health planners to provide more effective mobile health and telehealth interventions in rural areas, especially given that rural residents are equally likely to report that their providers utilize electronic health records [22].

Strengths of this study include the use of HINTS, which uses a scientifically rigorous probability-based sample and is nationally representative of the US adult population and oversamples for underrepresented populations. HINTS is meant to assess trends over time, with many core items having been collected repeatedly over the past 15 years; this allows for tracking of key metrics related to the HP2020 objectives. In addition, the weighting paradigm allows for weighted percent estimates reflective of the US population. Limitations include those associated with cross-sectional surveys, such as the inability to determine cause and effect. A second lies in the limitations of self-reported measures; that is, individuals may complete the survey with satisficing answers, not correctly recall information needed to answer the items, or leave items blank due to their sensitive nature. A final limitation is that response rates are lower than one would find in a prospective study of those actively engaged with the health care system.

Conclusions

The objectives and targets set forth by the HP2020 initiative for overall internet access and internet access via cellular network were met and, in the case of cellular network access, greatly surpassed. In addition, this study found that the digital divide still exists for many; however, the magnitude of these gaps is narrowing for several groups. However, internet access via traditional broadband delivery to the home began to decline before meeting the HP2020 target value. We believe that this is reflective of the rapidly changing technology landscape; that is, the objectives included in HP2020 were very specific to the means of accessing the internet available at the time the objectives were initially written, and the adoption of technology—and technology itself—evolved far more quickly than could have been anticipated. In creating objectives for HP2030, it could be beneficial to use wording broad enough to accommodate changes in modality for accessing the internet.

Acknowledgments

HINTS is funded by the NCI's Division of Cancer Control and Population Sciences with contract support to Westat, Inc, Rockville, MD, USA (contract number HHSN26120120000291).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Weighted multivariate logistic regression model of predictors of having internet access. Data from the National Cancer Institute's Health Information National Trends Survey administrations between 2003 and 2017 (n=30,150).

[DOCX File, 18KB - jmir_v21i6e13300_app1.docx]

Multimedia Appendix 2

Weighted multivariate logistic regression model of predictors of having broadband internet access among those who reported having internet access. Data from the National Cancer Institute's Health Information National Trends Survey administrations between 2003 and 2017 (n=20,150).

XSL•FO
RenderX

[DOCX File, 18KB - jmir_v21i6e13300_app2.docx]

References

1. Koh HK. A 2020 vision for healthy people. N Engl J Med 2010 May 6;362(18):1653-1656. [doi: 10.1056/NEJMp1001601] [Medline: 20445177]
2. Nelson DE, Kreps GL, Hesse BW, Croyle RT, Willis G, Arora NK, et al. The health information national trends survey (HINTS): development, design, and dissemination. J Health Commun 2004;9(5):443-60; discussion 81-4. [doi: 10.1080/10810730490504233] [Medline: 15513791]
3. Healthy People. 2018. Health Communication and Health Information Technology URL: https://www.healthypeople.gov/2020/topics-objectives/topic/health-communication-and-health-information-technology/objectives [accessed 2018-11-03]
4. Hesse BW, Gaysynsky A, Ottenbacher A, Moser RP, Blake KD, Chou WS, et al. Meeting the healthy people 2020 goals: using the health information national trends survey to monitor progress on health communication objectives. J Health Commun 2014;19(12):1497-1509 [FREE Full text] [doi: 10.1080/10810730.2014.954084] [Medline: 25491584]
5. Serrano KJ, Thai CL, Greenberg AJ, Blake KD, Moser RP, Hesse BW. Progress on broadband access to the internet and use of mobile devices in the United States. Public Health Rep 2017;132(1):27-31 [FREE Full text] [doi: 10.1177/0033354916679365] [Medline: 28005473]
6. Westat. Health Information National Trends Survey. 2015. Health Information National Trends Survey 4 (HINTS 4): Cycle 1 Methodology Report URL: http://www.hints.cancer.gov
7. Blake KD, Portnoy DB, Kaufman AR, Lin CT, Lo SC, Backlund E, et al. Rationale, procedures, and response rates for the 2015 administration of NCI's health information national trends survey: HINTS-FDA 2015. J Health Commun 2016;21(12):1269-1275 [FREE Full text] [doi: 10.1080/10810730.2016.1242672] [Medline: 27892827]
8. Finney RL, Davis T, Beckjord EB, Blake K, Moser RP, Hesse BW. Picking up the pace: changes in method and frame for the health information national trends survey (2011-2014). J Health Commun 2012;17(8):979-989 [FREE Full text] [doi: 10.1080/10810730.2012.700998] [Medline: 23020763]
9. 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 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
10. Brodie M, Flournoy RE, Altman DE, Blendon RJ, Benson JM, Rosenbaum MD. Health information, the internet, and the digital divide. Health Aff (Millwood) 2000;19(6):255-265. [doi: 10.1377/hlthaff.19.6.255] [Medline: 11192412]
11. Rains SA. Health at high speed: broadband internet access, health communication, and the digital divide. Commun Res 2008 Apr 14;35(3):283-297. [doi: 10.1177/0093650208315958]
12. United States Department of Agriculture: Economic Research Service. 2016. Rural-Urban Continuum Codes URL: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/ [accessed 2018-11-03]
13. Health Information National Trends Survey. URL: https://hints.cancer.gov/
14. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. Info Systems J 2018 Jan 29;28(6):995-1019 (forthcoming). [doi: 10.1111/isj.12179]
15. Tsetsi E, Rains SA. Smartphone internet access and use: extending the digital divide and usage gap. Mob Media Commun 2017 Jun 13;5(3):239-255. [doi: 10.1177/2050157917708329]
16. Christopher GM. Use of health information technology among racial and ethnic underserved communities. Perspect Health Inf Manag 2011;8:1f [FREE Full text] [Medline: 21307989]
17. Silva S, Badasyan N, Busby M. Diversity and digital divide: using the national broadband map to identify the non-adopters of broadband. Telecomm Policy 2018 Jun;42(5):361-373. [doi: 10.1016/j.telpol.2018.02.008]
18. Anderson-Lewis C, Darville G, Mercado RE, Howell S, di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
19. Bartikowski B, Laroche M, Jamal A, Yang Z. The type-of-internet-access digital divide and the well-being of ethnic minority and majority consumers: a multi-country investigation. J Bus Res 2018 Jan;82:373-380. [doi: 10.1016/j.jbusres.2017.05.033]
20. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018 Aug 1;25(8):1080-1088. [doi: 10.1093/jamia/ocy052] [Medline: 29788380]
21. Federal Communications Commission. 2018. 2018 Broadband Deployment Report URL: https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2018-broadband-deployment-report
22. Greenberg AJ, Haney D, Blake KD, Moser RP, Hesse BW. Differences in access to and use of electronic personal health information between rural and urban residents in the United States. J Rural Health 2018 Feb;34(Suppl 1):s30-s38 [FREE Full text] [doi: 10.1111/jrh.12228] [Medline: 28075508]

Abbreviations

DSL: digital subscriber line
FiOS: fiber optic service

XSL•FO
RenderX

HC/HIT: Health Communication and Health Information Technology
HINTS: Health Information National Trends Survey
HP2020: Healthy People 2020
NCI: National Cancer Institute
ODPHP: Office of Disease Prevention and Health Promotion
OR: odds ratio
RUCCs: Rural-Urban Continuum Codes

Edited by G Eysenbach; submitted 03.01.19; peer-reviewed by E Beckjord, K Colorafi, J Hu, V Foufi; comments to author 27.04.19; revised version received 04.05.19; accepted 05.05.19; published 24.06.19.

Please cite as:
Greenberg-Worisek AJ, Kurani S, Finney Rutten LJ, Blake KD, Moser RP, Hesse BW
Tracking Healthy People 2020 Internet, Broadband, and Mobile Device Access Goals: An Update Using Data From the Health Information National Trends Survey
J Med Internet Res 2019;21(6):e13300
URL: http://www.jmir.org/2019/6/e13300/
doi:10.2196/13300
PMID:31237238

©Alexandra J Greenberg-Worisek, Shaheen Kurani, Lila J Finney Rutten, Kelly D Blake, Richard P Moser, Bradford W Hesse. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 24.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Rapid Analysis of Diagnostic and Antimicrobial Patterns in R (RadaR): Interactive Open-Source Software App for Infection Management and Antimicrobial Stewardship

Christian Friedemann Luz[1], MSc, MD; Matthijs S Berends[1,2], MSc; Jan-Willem H Dik[1], PhD; Mariëtte Lokate[1], PhD; Céline Pulcini[3,4], MD, PhD; Corinna Glasner[1], MSc, PhD; Bhanu Sinha[1], MD, PhD

[1]Department of Medical Microbiology and Infection Prevention, University Medical Center Groningen, University of Groningen, Groningen, Netherlands

[2]Certe Medical Diagnostics and Advice, Groningen, Netherlands

[3]APEMAC, Université de Lorraine, Nancy, France

[4]Infectious Diseases Department, CHRU-Nancy, Université de Lorraine, Nancy, France

Corresponding Author:
Christian Friedemann Luz, MSc, MD
Department of Medical Microbiology and Infection Prevention
University Medical Center Groningen
University of Groningen
Hanzeplein 1
Groningen, 9713 GZ
Netherlands
Phone: 31 50 361 3480
Fax: 31 50 36 19105
Email: c.f.luz@umcg.nl

Abstract

Background: Analyzing process and outcome measures for all patients diagnosed with an infection in a hospital, including those suspected of having an infection, requires not only processing of large datasets but also accounting for numerous patient parameters and guidelines. Substantial technical expertise is required to conduct such rapid, reproducible, and adaptable analyses; however, such analyses can yield valuable insights for infection management and antimicrobial stewardship (AMS) teams.

Objective: The aim of this study was to present the design, development, and testing of RadaR (Rapid analysis of diagnostic and antimicrobial patterns in R), a software app for infection management, and to ascertain whether RadaR can facilitate user-friendly, intuitive, and interactive analyses of large datasets in the absence of prior in-depth software or programming knowledge.

Methods: RadaR was built in the open-source programming language R, using Shiny, an additional package to implement Web-app frameworks in R. It was developed in the context of a 1339-bed academic tertiary referral hospital to handle data of more than 180,000 admissions.

Results: RadaR enabled visualization of analytical graphs and statistical summaries in a rapid and interactive manner. It allowed users to filter patient groups by 17 different criteria and investigate antimicrobial use, microbiological diagnostic use and results including antimicrobial resistance, and outcome in length of stay. Furthermore, with RadaR, results can be stratified and grouped to compare defined patient groups on the basis of individual patient features.

Conclusions: AMS teams can use RadaR to identify areas within their institutions that might benefit from increased support and targeted interventions. It can be used for the assessment of diagnostic and therapeutic procedures and for visualizing and communicating analyses. RadaR demonstrated the feasibility of developing software tools for use in infection management and for AMS teams in an open-source approach, thus making it free to use and adaptable to different settings.

(J Med Internet Res 2019;21(6):e12843) doi:10.2196/12843

KEYWORDS

antimicrobial stewardship; software; hospital records; data visualization; infection, medical informatics applications

XSL•FO
RenderX

Introduction

Background

With antimicrobial resistance (AMR) on the rise, efforts are being made worldwide to focus on the preservation of antimicrobials as a precious nonrenewable resource. Infection management in the form of antimicrobial stewardship (AMS) programs has emerged as an effective solution to address this global health problem in hospitals. AMS programs are defined as "a coherent set of actions which promote using antimicrobials responsibly" [1]. Stewardship interventions and activities focus on individual patients (personalized medicine and consulting) as well as patient groups or clinical syndromes (guidelines, protocols, information technology infrastructure, and clinical decision support systems) while prioritizing improvement in quality of care and patient safety for any intervention. The appropriate use of antimicrobials based on accurate and timely diagnostics is integral for the successful management of infections. In doing so, the diagnostics contribute to efforts in minimizing AMR by optimizing the use of antimicrobials.

AMS setups in hospitals are often heterogeneous, but audit and feedback to assess the goals are essential parts of most programs, and they are included in international guidelines and reviews [2-7]. Important data for AMS programs include, for example, days of therapy (DOT), daily defined doses (DDD), admission dates, length of stay (LOS), and adherence to local or national diagnostic, therapeutic, or infection management guidelines [1]. Clinical outcomes, quality of care, or consumption of hospital resources can be measured, for example, using mortality data or surrogate parameters such as LOS. The collection of these data is facilitated by electronic health records (EHRs) and administrative local databases. Notably, administrative data have also been shown to be a reliable source for assessing clinical outcomes [8].

EHRs usually offer quick insights into useful infection management data on the individual patient level. However, easy access to analyze patient groups (eg, stratified by departments or wards, specific antimicrobials, or diagnostic procedures used) is difficult to implement in daily practice. It is even more challenging to rapidly analyze larger patient populations (eg, spread over multiple specialties) even though this information might be available. Nevertheless, this is vital for meaningful analysis, including possible confounders and pattern recognition across different populations. Moreover, when aggregated data are available, it is often not possible to trace individual patients, and analyses lack the ability to be further adjusted or stratified.

AMS teams are multidisciplinary, and they act beyond the borders of single specialties [9]. They are usually understaffed, with limited data analysis support [10,11]. Therefore, they need user-friendly and time-saving data analysis resources, without the need for profound technical expertise once the system is set up. Aggregating and linking data of antimicrobial use, guideline adherence, and clinical outcomes at the institutional level can build the basis for important insights for these teams. These could be used to identify areas within hospitals that might benefit most from supportive AMS interventions (eg, subspecialties with lower guideline adherence or unusual patterns of antimicrobial use). Moreover, feedback from these data could help physicians better understand their patient population as a whole; in addition, hospital administration could allocate resources in a more targeted fashion.

Furthermore, aggregated data and simultaneous analysis of multiple areas (eg, use of diagnostics and antimicrobials) present an extensive insight into large patient populations. This also enables the development of comprehensive and multidisciplinary approaches of infection management, combining diagnostic and therapeutic perspectives [1,9,12]. Unfortunately, these kinds of analyses still require substantial statistical knowledge and software skills, and it is time consuming when performed.

Technology, data science, and software app development can bring solutions to complex data handling problems such as those described above. Software app development for medical and epidemiological (research) questions has found many important answers during recent years. For example, software apps at hospital emergency departments (EDs) in the form of a dashboard have been shown to improve efficiency and quality of care for patients requiring emergency admission to hospital [13]. These software apps are used to communicate clearly defined clinical problems, such as mortality ratio, number of cardiac arrests, or readmission rate to the EDs. This has led to a decreased LOS and mortality at the EDs. Others used similar approaches to rapidly and interactively display geographical locations of tuberculosis cases without the need of technical expertise improving the understanding of transmission and detection [14]. Furthermore, data-driven fields such as genomics are front runners in developing new, innovative software apps to handle large datasets, in close collaboration with bioinformatics [15]

It is important to note that all of these abovementioned software apps have been created in an open-source approach. This means that the underlying source code can be easily shared, easily modified, and freely distributed through open repositories, such as GitHub [16], taking open-source software license obligations into account. This facilitates collaboration, quality control through code review, and easy adaptation to many different settings and information technology systems, and this supports the use of advanced data visualization for users with minimal experience in programming and little or no budget for professional database engineers [15].

In the field of medical microbiology, different approaches have been described to interactively work with microbiological diagnostics data and EHRs: electronic antibiograms, centralized resistance analysis, EHR data mining, and clinical decision support systems for AMS are great examples for innovation in the field [17-19]. However, a full open-source approach for software apps working with combined antimicrobials use and diagnostic data of individual patients on the hospital level in the field of infection management is still lacking.

Objectives

We followed principles of open knowledge [20] to address the need for an interactive, easy-to-use software app that allows users to investigate antimicrobial use, microbiological diagnostic use, and patient outcomes at an institutional (hospital) level.

XSL•FO
RenderX

We developed an open-source, Web-based software app—Rapid analysis of diagnostic and antimicrobial patterns in R (RadaR) that can be used for AMS and infection management. This free software app can be run on regular computers or implemented on local or Web-based servers to be accessed through standard Web browsers. The focus user group of this software app is health care professionals involved in AMS (eg, infectious disease specialists, clinical microbiologists, and pharmacists). Although some technical expertise (basic R knowledge) is needed for installation and implementation, the use of RadaR follows usual Web browser user experiences. RadaR enables rapid and reproducible data analysis without extensive previous analysis expertise in a graphically appealing way while being adaptable to different settings. RadaR's analyses are based on datasets of individual patients. Therefore, aggregated results can also be stripped down, and additional patient features can be investigated. With this software app, we aim at supporting data-driven hospital insights and decision making for actors in the field of AMS in a free, transparent, and reproducible way.

Methods

For the development of software in an open-source environment, we used the open-source programming language R in conjunction with RStudio version 1.1.463 (RStudio, Inc) [21], an open-source integrated desktop environment for R [22]. Both R and RStudio are free of charge, and they need to be installed for the development and implementation of RadaR. To build RadaR as a Web-based software app, we used the Shiny package for R [23]. Shiny allows R users to build interactive Web apps without extensive knowledge in Web design and its programming languages. The Web apps can be run and hosted on the Web for free [24], as well as on local or cloud-based servers or on personal computers.

The functionality of R can be easily extended by installing additional packages. All packages used for the development of RadaR are listed in Table 1. RadaR is developed in an open-source environment and licensed under GNU General Public License v2.0 [25], giving options to change, modify, and adapt RadaR to both personal and commercial users' needs while requiring the need to document code changes [25].

RadaR's calculations and data aggregation are done reactively on the basis of the selection of the user. Single observations on the patient level build the basis for any calculation. RadaR uses common CSV files as input. A total of 3 different data sources are read in RadaR for admission, antimicrobial, and microbiological data, which are merged and transformed upon start. A patient number or study number is used as a unique identifier. All antimicrobial and microbiological data are checked to ascertain whether they fall in the interval of admission dates.

Table 1. Required R packages for RadaR.

R package	Minimal version
AMR	0.5.0
data.table	1.11.6
DT	0.4.0
ggridges	0.5.0
lubridate	1.7.4
plotly	4.8.0
qicharts2	0.5.1
rintrojs	0.2.0
shiny	1.1.0
shinyBS	0.61
shinycssloaders	0.2.0
shinydashboard	0.7.0
shinyjs	1.0.0
shinyWidgets	0.4.3
survival	2.42-6
survminer	0.4.3
tidyverse	1.2.1
viridis	0.5.1
zoo	1.8-3

XSL•FO
RenderX

Table 2. Input variables for RadaR.

Variable	Detail
Admission data	
adm_end_date	Discharge date[a]
adm_id	Admission ID
adm_route	Origin
adm_start_date	Admission date[a]
birth_date	Birth date[a]
death_during_adm	In-hospital death (TRUE/FALSE)
gender	Gender
id	Patient ID or study ID
specialty	General specialty (internal medicine, surgery, and other)
sub_specialty	Subspecialty
Antimicrobial data	
ab_route	Administration route
ab_start_date	Start of antimicrobial[a]
ab_stop_date	Stop of antimicrobial[a]
atc_code	Fifth level of the World Health Organization Anatomical Therapeutic Chemical (WHO ATC) classification system[b]
ddd_per_day	Defined daily dose of antimicrobial according to WHO ATC classification system per day[b]
id	Patient ID or study ID
Microbiological data	
antimicrobial susceptibility testing	Several columns of tested antimicrobial agents (eg, amoxicillin, ciprofloxacin) with resistance results (R/I/S)
id	Patient ID or study ID
material	Test material
mo	Microbial ID (if test=positive)[c]
specialty	Ordering specialty
test_date	Test date[a]

[a]YYYY-MM-DD.

[b]As available on the website [31].

[c]As defined by the AMR package for R [30].

The input data should be structured in a dataset format, where each variable is 1 column and each observation is 1 row. This follows the concept of "tidy data," as defined by Hadley Wickham [26]. Table 2 displays the set of variables underlying RadaR's functionality. In our setting for the development of RadaR, these variables originated from 3 different data sources: administrative data from the hospital data warehouse, microbiological data from the laboratory information system, and antimicrobial prescription data from the computerized prescriber order entry system. The data preparation and cleaning process are very specific for each data source, dependent on local data standards, and difficult to generalize. Therefore, Table 2 represents the final variables and formats for the analysis and use with RadaR, referring to the "tidy data" concept above and to the tidyverse R package collection for the preparation process [26,27]. Additional variables are calculated and transformed using the packages lubridate and zoo for time points and intervals, and AMR for antimicrobial (group) names, microbial isolate names, first isolate identification, and resistance analysis [28-30]. Microbiological resistance is calculated per antimicrobial substance or as coresistance if more than 1 substance is selected.

RadaR can be used for graphical exploratory data analysis. Differences in LOS are displayed by a Kaplan-Meier curve in conjunction with a log-rank test, using the survminer package [32]. Time trends for number of admissions, antimicrobial consumption, and resistance counts per year, quarter, or month, are visualized in run charts using the qicharts2 package [33].

XSL•FO
RenderX

Nonrandom variation in these run charts is tested using Anhøj's rules [34].

RadaR has been developed in macOS High Sierra (1.4 GHz, 4 GB RAM), and it was successfully tested in Windows 7 (3.2 GHz, 8 GB RAM) and Linux (Ubuntu 16.04.4 LTS, 3.4 GHz, 12 GB RAM). A running example version has been deployed to shinyapps.io, a publicly available Web hosting service for R Shiny apps [35]. The entire source code of RadaR is freely accessible on GitHub [36]. We intend to integrate suggestions and feedback coming from its users and the R community. RadaR was developed using data of patients admitted to the University Medical Center Groningen, Groningen, the Netherlands. Data were collected retrospectively, and permission was granted by the ethical committee (METc 2014/530). RadaR can be used locally in protected environments or hosted on the Web, provided appropriate measures have been taken to guarantee data protection, depending on national regulations.

Results

Overview

We have developed RadaR, a Web-based software app providing an intuitive platform for rapid analysis of large datasets containing information about patients' admission, antimicrobial use, and results of microbiological diagnostic tests. This software app can help users (ie, AMS team members) find answers to questions, such as "What are the most commonly used antimicrobials at an institution/specialty/department and have they changed over time?," "Were adequate microbiological diagnostics performed at the start of antimicrobial treatments?," "What are the most frequent microorganisms found and their resistance patterns in different departments?," and "Can we identify priority areas within a hospital where antimicrobial or microbiological diagnostic use has the largest room for improvement?"

Application Design

RadaR is designed in the form of a Web browser–based dashboard that most users are familiar with from typical websites and Web-based tools (see Figure 1). The basis of RadaR's functionality is filtering datasets and producing analytical graphs according to selection criteria defined by the user. Any calculations and data aggregation are based on single observations of individual patients. To identify and analyze groups of patients, 17 different selection criteria can be found in the sidebar (Table 3). The output of RadaR is grouped into 4 panels (patient, antimicrobials, diagnostics, and outcome) that each comprise 3 to 4 output boxes displaying the results (see Multimedia Appendix 1).

All output is based on the selection criteria defined by the user in the sidebar. Each new selection and any change need to be confirmed by clicking the confirm selection button (see Figure 1). Users can navigate among the different analysis panels by clicking the respective button.

Figure 1. Application design.

XSL•FO
RenderX

Table 3. Selection criteria in sidebar.

Tab name and criteria	Functionality
Antimicrobials	
Start of antimicrobials (in relation to start of admission)	Select patients starting treatment in a defined time period
Minimum duration of treatment (days)—all antimicrobials	Select patients with a minimum treatment duration
Minimum duration of prescription (days)—single antimicrobial	Define the minimum duration of a prescription for any selected antimicrobial
Administration route	Intravenous or oral
First antimicrobial only	Filter patients for first prescribed antimicrobial only or any (on the basis of all other selection criteria)
Groups of antimicrobials	Fourth level of the World Health Organization Anatomical Therapeutic Chemical (WHO ATC) classification system[a]
Antimicrobials	Fifth level of the WHO ATC classification system[a]
Patients	
Gender	Female or male
Age	As available in the data
Year	
Year	Years available in the data
Specialty	
Specialty	Internal medicine, surgery, or other
Minimum number of patients per subspecialty	0, 10, 100, 1000, or 10,000
Include only this subspecialty	All other subspecialties will be excluded
Exclude subspecialty	Define single subspecialties to be excluded
Origin	
Origin at admission	As available in the data
Diagnostics	
Type of diagnostics	Blood culture or urine culture test
Days to first test (in relation to start of antimicrobials)	Define time period for tests to be performed in

[a]As available on the website [31].

Results are shown in bar charts, density plots, run charts, a bubble plot, and a Kaplan-Meier curve for LOS in hospital. Each panel further displays a table summarizing the respective data analyses. All output boxes and their content are described in Table 4. Most output boxes include modification options that can be identified by small gear icons (see Figure 1). These clickable icons allow for further specification of the generated plots and tables. Users can compare different groups (eg, antimicrobial use by antimicrobial agent, resistance patterns per isolate, or LOS by specialty) or modify the plots (eg, switch from count to proportion, change the chart type, or show or hide the legend). Plots and tables can be downloaded through download buttons as PNG files for plots and CSV, Excel, or PDF files for tables.

Finally, 2 datasets (antimicrobial/admission data and microbiological data) of the user-defined selection can be downloaded from the sidebar menu in a CSV-file format for further analysis (eg, retrieving a list of patient numbers of the selected patient group).

XSL•FO
RenderX

Table 4. Output boxes for analysis results.

Output panel and output box	Output type	Content	Modification options
Patients			
Subspecialties in selection	Bubble chart[a]	Patients per subspecialty	Show top 10 by number of patients
Subspecialties—table	Table	Total number of patients and per subspecialty	—[b]
Patient age	Density plot (distribution)	Age distribution in selection	Group by gender
Number of admissions	Run chart	Count of admissions per time period	Per year, per quarter, or per month
Antimicrobials			
Antimicrobials	Bar chart	(Group of) antimicrobials sorted by prescription, DDD[c], or DOT[d]	Single antimicrobials or groups; select prescription count, DDD, or DOT per 100 bed days
DDD	Run chart	DDD per 100 bed days per group and per time period	Group by none, specialty, subspecialty, and origin; per year, per quarter, and per month
DOT	Run chart	DOT per 100 bed days per group and per time period	Group by none, specialty, subspecialty, and origin; per year, per quarter, and per month
DDD/DOT table	Table	Summary of DDD/DOT per 100 bed days per group	DDD or DOT per 100 bed days; group by antimicrobial (group), year, specialty, subspecialty, and origin
Diagnostics			
Diagnostics in selected patients	Bar chart	Diagnostics taken versus not taken in specified timespan	Count or proportion; per year, quarter, or month
Timing of selected diagnostics	Bar chart	Time of diagnostics performed in days after start of treatment	—[b]
Diagnostics in relation	Bar chart	Absolute difference from average proportion of selected diagnostics performed	Group by antimicrobial (group), year, specialty, subspecialty, and origin
Table—proportion performed	Table	Summary of proportion of diagnostics performed	Group by antimicrobial (group), year, specialty, subspecialty, and origin
First isolates in selected diagnostics	Bar chart	First isolates of microorganisms sorted by frequency	Group by antimicrobial (group), year, specialty, subspecialty, and origin; zoom to select more or less isolates shown in graph
First isolates—table	Table	Frequency table of first isolates	Group by year, specialty, subspecialty, and origin
Resistance analysis	Bar chart	Count or proportion of resistance or coresistance to selected antimicrobials in selected isolates in “R,” “S,” and “I” categories	Select isolates; select antimicrobials; group by year, specialty, subspecialty, and origin; select count or proportion
Resistance—over time	Run chart	Count of resistance or coresistance to selected antimicrobials in selected isolates in “R,” “S,” and “I” categories over time	Select isolates; select antimicrobials; per year, per quarter, or per month
Table	Table	Count or proportion of resistance or coresistance to selected antimicrobials in selected isolates in “R,” “S,” and “I” categories	Group by year, month, quarter, specialty, subspecialty, and origin; select isolates; select antimicrobials; select count or proportion
Outcome			
Length of stay	Density plot or histogram	Distribution of length of stay per group	Group by all, gender, year, antimicrobial (group), diagnostics performed, specialty, subspecialty, and origin; show histogram; show legend; spread out to remove overlaps
Length of stay—Kaplan-Meier	Kaplan-Meier curve	Kaplan-Meier curve per group	Groups shown as selected in the length-of-stay box
Length of stay—table	Table	Summary of length of stay per group	Group by gender, year, antimicrobial (group), diagnostics performed, specialty, subspecialty, and origin

XSL•FO
RenderX

[a]Interactive plot showing additional information when hovering over plot.

[b]Not applicable.

[c]DDD: defined daily doses.

[d]DOT: days of therapy.

Development Process

RadaR has been developed in close contact with the AMS team and senior consulting specialists at the University Medical Center Groningen, Groningen, the Netherlands, to meet the needs and requirements of this user group. Subsequently, all members of the European Society of Clinical Microbiology and Infectious Diseases Study Group for Antimicrobial Stewardship (ESGAP) were asked to evaluate and test the software app through a running Web-based example of RadaR and by filling out a Web-based survey. The ESGAP comprises around 200 members from more than 30 countries worldwide. A total of 12 members from 9 different countries took part in the evaluation. This yielded important information on user experiences with the software app, which in turn led to further improvements that are reflected in the version we presented in this report. In a next phase, RadaR will be tested in different settings of ESGAP members and other interested partners using locally available data (eg, an 837-bed tertiary care hospital in the Netherlands and a 750-bed tertiary care hospital in Greece).

Workflow

RadaR was developed and tested with a dataset of all patients admitted to our institution, a 1339-bed academic tertiary referral hospital, within the years of 2009 to 2016, comprising over 180,000 admissions. For simulation purposes and Web-based user testing, we have created a test dataset of 60,000 simulated patients. This sample dataset allows testing of RadaR's functionality, but it does not produce meaningful results.

A typical example workflow with RadaR comprises 6 steps (with examples from the test dataset). They are listed below:

1. Define the selection: For example, patients receiving intravenous second- or third-generation cephalosporins as first treatment for at least 2 days, starting within the first 2 days of hospital admission from any specialty in all years in the dataset.
2. Patients' panel: Identify the total number of patients and the subspecialties with the highest number of included patients (eg, 537 patients selected in total, with 97 patients from internal medicine). Investigate patients' gender and age distribution.
3. Antimicrobials panel: Identify the total use of the initial cefuroxime treatment in DDD and DOT per 100 bed days (eg, 4.51 and 1.5, respectively). Stratify the results by subspecialty and identify the highest number of DDD and DOT per 100 bed days (eg, highest use by DDD and DOT in internal medicine).
4. Diagnostics panel: Check if the selected microbiological diagnostic test (eg, blood culture test) has been performed on the same day as the start of the treatment (defined in the sidebar). Investigate the proportion of tests performed over the years and investigate which subspecialty performs best compared with others (eg, Pediatrics). Check which microorganisms (as first isolates) were found in the selected diagnostic specimens (the most common isolate: *Escherichia coli*). Investigate the proportion of isolates resistant to cefuroxime (8.9%) and analyze the trend over time.
5. Outcome panel: Check for patterns of differences in LOS in the defined patient group by subspecialties or performed diagnostics (eg, highest mean LOS of 7.8 days in Surgery).
6. Refine the selection: Investigate a subgroup of the original selection. For example, select only the top 3 subspecialties by number of patients and repeat step 2 to 5.

Customization

For setting up RadaR in a new environment after data preparation, users only need to perform the following 4 steps:

1. Downloading R and RStudio [21,22], which are free to use and open-source software
2. Download or copy and paste RadaR's source code [36] into 3 files in RStudio—global.R, server.R, and ui.R
3. In global.R, manually edit the paths for the prepared datasets to be imported into RadaR
4. Run the app in RStudio with the calling the function runApp() in the console or by clicking the green run app button. This will download and install the required R packages needed for the app if they have not been installed previously, and this will create the final dataset for analysis. The RadaR interface will open in the RStudio viewer pane or in a new window of the standard browser of the user's operating system.

RadaR's appearance has been customized using a cascading style sheets (CSS) script [37] that is loaded into the app upon its start. This script needs to be saved into a subdirectory of the directory of the 3 main files (global.R, server.R, and ui.R) called "www." We recommend RStudio's project function to create a single project for RadaR and to store all information in this project directory. Users with experience in using CSS can fully alter RadaR's design by changing the underlying CSS script.

Discussion

Principal Findings

We have developed a Web-based software app for rapid analysis of diagnostic and antimicrobial patterns that can support AMS teams to tailor their interventions. It has been designed to enhance communication of relevant findings while being easy to use. This also applies to users without extensive prior software skills, as it follows usual Web browser user experiences. Moreover, it has been developed using open-source software. It is therefore free to use and accessible for download. In our experience, this system can be adapted to new settings within 1 day, when the required data (Table 2) are available.

Commercial software for infection management is available (eg, Epic Antimicrobial Stewardship Module, TREAT Steward).

XSL•FO
RenderX

These offer extensive options for filtering, analyzing, and visualizing EHRs with real-time connections to hospital data infrastructures and have been shown to be useful in clinical practice [38]. However, it is difficult to compare functionalities of these tools because of their non–open-source nature. This fact, along with the required budget to purchase the software, drastically limits their use. We are convinced that transparent software development can support the adoption of data-driven developments while enhancing optimal quality of care and patient safety, which is crucial in the light of new data-driven developments of using EHRs [39,40].

The global nature of infections further calls to develop software tools applicable in resource-limited settings [41]. Open-source approaches for data analysis, such as RadaR, have advantages over traditional methods, such as Excel or SPSS. Hughes et al described those in their report of a software app for RNA-sequencing data analysis [15]. They highlight aspects that were also fundamental for the development of RadaR. First, R allows transparent, reproducible, and sustainable data analysis through scripts that can easily be shared and changed. This can build the basis for collaboration, and this enforces the spirit of open science (also through the strong collaborative R community on the Web). Second, R is open source and free to use; therefore, it also enables use in resource-limited settings. Finally, Shiny empowers users to interact with the data, making even very large datasets quickly interpretable.

Innovative approaches used in supporting infection management by leveraging EHRs are being investigated [17-19]. Reporting on AMR, antimicrobial use, and hospital infections (eg, for quality assurance) is well established, but it is important to integrate these data sources in an approach that allows detailed filtering options on all input. Merely looking at antimicrobial use alone or comparing aggregated results (eg, total amount of a specific antimicrobial substance per hospital correlated with the total count of a resistant isolate) will result in loss of information or even misleading interpretation. Detailed data and calculations on the basis of each individual patient are crucial to draw informed conclusions. Unfortunately, the abovementioned infection management approaches [17-19] either depend on additional commercial software for data visualization or the source code is not openly available. We want to encourage others to turn toward available open-source software solutions, such as R, for an increased potential of collaboration and transparency. However, their strength is the connection to real-time data flows. This enables the prospective use and increases their usability for daily clinical practice. RadaR is currently still limited to retrospective data analysis because of a changing hospital data infrastructure in our setting. Technically, it is feasible to connect R-based software apps such as RadaR to real-time hospital data infrastructures running with clinical data standards [42]. For a start, access to static data extraction is often easier and faster to achieve. RadaR can be used to advocate the use of data visualization tools and improved accessibility of hospital data sources. Until connection to real-time hospital data is established, RadaR can support users as a stand-alone option for retrospective data analysis in infection management. Next steps will involve testing in multiple settings and forming a user and research group to continue and expand the use of open-source technology and open science principles in infection management.

Conclusions

RadaR demonstrates the feasibility of developing software tools for infection management and AMS teams in an open-source approach, making it free to use, share, or modify according to various needs in different settings. RadaR has the potential to be a highly useful tool for infection management and AMS in daily practice.

Acknowledgments

The authors would like to thank the ESGAP executive committee for supporting the evaluation of RadaR in the ESGAP study group and all its members for their valuable input, suggestions, and comments. Furthermore, the authors wish to thank Igor van der Weide, Jan Arends, and Prashant Nannan Panday for their great support in obtaining required data at our institution that built the basis for the development of RadaR. The authors also thank the online R community as well as the valuable comments, suggestions, and input from reviewers that they have received to improve RadaR.

RadaR was developed as part of a project funded by the European Commission Horizon 2020 Framework Marie Skłodowska-Curie Actions (grant agreement number: 713660-PRONKJEWAIL-H2020-MSCA-COFUND-2015).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot of RadaR.

[MOV File, 6MB - jmir_v21i6e12843_app1.mov]

References

XSL•FO
RenderX

1. Dyar OJ, Huttner B, Schouten J, Pulcini C, ESGAP (ESCMID Study Group for Antimicrobial stewardshiP). What is antimicrobial stewardship? Clin Microbiol Infect 2017 Nov;23(11):793-798 [FREE Full text] [doi: 10.1016/j.cmi.2017.08.026] [Medline: 28882725]
2. Barlam TF, Cosgrove SE, Abbo LM, MacDougall C, Schuetz AN, Septimus EJ, et al. Implementing an Antibiotic Stewardship Program: guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. Clin Infect Dis 2016 May 15;62(10):e51-e77. [doi: 10.1093/cid/ciw118] [Medline: 27080992]
3. Davey P, Marwick CA, Scott CL, Charani E, McNeil K, Brown E, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients. Cochrane Database Syst Rev 2017 Dec 9;2:CD003543. [doi: 10.1002/14651858.CD003543.pub4] [Medline: 28178770]
4. National Institute for Health and Care Excellence. Antimicrobial stewardship: systems and processes for effective antimicrobial medicine use URL: https://www.nice.org.uk/Guidance/NG15 [accessed 2018-11-16] [WebCite Cache ID 73ybLqgCI]
5. Schuts EC, Hulscher ME, Mouton JW, Verduin CM, Stuart JW, Overdiek HW, et al. Current evidence on hospital antimicrobial stewardship objectives: a systematic review and meta-analysis. Lancet Infect Dis 2016 Jul;16(7):847-856. [doi: 10.1016/S1473-3099(16)00065-7] [Medline: 26947617]
6. Stichting Werkgroep Antibiotica Beleid. 2016. SWAB Guidelines for Antimicrobial Stewardship 2016 URL: http://www.swab.nl/swab/cms3.nsf/uploads/FAF701FBF4FCBDD6C12580EC0037D292/$FILE/20161228%20SWAB%20Guidelines%20for%20Antimicrobial%20Stewardship%20-definitief.pdf [accessed 2018-11-16] [WebCite Cache ID 73ycxKUfO]
7. Pulcini C, Binda F, Lamkang AS, Trett A, Charani E, Goff DA, et al. Developing core elements and checklist items for global hospital antimicrobial stewardship programmes: a consensus approach. Clin Microbiol Infect 2018 Apr 3:20-25. [doi: 10.1016/j.cmi.2018.03.033] [Medline: 29625170]
8. Sarkies MN, Bowles K, Skinner EH, Mitchell D, Haas R, Ho M, et al. Data collection methods in health services research: hospital length of stay and discharge destination. Appl Clin Inform 2015;6(1):96-109 [FREE Full text] [doi: 10.4338/ACI-2014-10-RA-0097] [Medline: 25848416]
9. British Society for Antimicrobial Chemotherapy. Birmingham, United Kingdom: British Society for Antimicrobial Chemotherapy; 2018. Antimicrobial stewardship: From Principal to Practice URL: http://bsac.org.uk/antimicrobial-stewardship-from-principles-to-practice-e-book/ [accessed 2018-11-16] [WebCite Cache ID 73ydEbLNu]
10. Pulcini C, Morel CM, Tacconelli E, Beovic B, de With K, Goossens H, et al. Human resources estimates and funding for antibiotic stewardship teams are urgently needed. Clin Microbiol Infect 2017 Nov;23(11):785-787 [FREE Full text] [doi: 10.1016/j.cmi.2017.07.013] [Medline: 28778544]
11. Howard P, Pulcini C, Levy Hara G, West RM, Gould IM, Harbarth S, ESCMID Study Group for Antimicrobial Policies (ESGAP), ISC Group on Antimicrobial Stewardship. An international cross-sectional survey of antimicrobial stewardship programmes in hospitals. J Antimicrob Chemother 2015 Apr;70(4):1245-1255. [doi: 10.1093/jac/dku497] [Medline: 25527272]
12. Dik JH, Poelman R, Friedrich AW, Panday PN, Lo-Ten-Foe JR, van Assen S, et al. An integrated stewardship model: antimicrobial, infection prevention and diagnostic (AID). Future Microbiol 2016;11(1):93-102 [FREE Full text] [doi: 10.2217/fmb.15.99] [Medline: 26323589]
13. Staib A, Sullivan C, Jones M, Griffin B, Bell A, Scott I. The ED-inpatient dashboard: Uniting emergency and inpatient clinicians to improve the efficiency and quality of care for patients requiring emergency admission to hospital. Emerg Med Australas 2017 Jun;29(3):363-366. [doi: 10.1111/1742-6723.12661] [Medline: 27592365]
14. Smith CM, Hayward AC. DotMapper: an open source tool for creating interactive disease point maps. BMC Infect Dis 2016 Apr 12;16:145 [FREE Full text] [doi: 10.1186/s12879-016-1475-5] [Medline: 27066780]
15. Hughes LD, Lewis SA, Hughes ME. ExpressionDB: an open source platform for distributing genome-scale datasets. PLoS One 2017;12(11):e0187457 [FREE Full text] [doi: 10.1371/journal.pone.0187457] [Medline: 29095940]
16. GitHub. URL: https://github.com [accessed 2018-11-16] [WebCite Cache ID 73yb4rqQT]
17. Simpao AF, Ahumada LM, Larru Martinez B, Cardenas AM, Metjian TA, Sullivan KV, et al. Design and implementation of a visual analytics electronic antibiogram within an electronic health record system at a tertiary pediatric hospital. Appl Clin Inform 2018 Dec;9(1):37-45. [doi: 10.1055/s-0037-1615787] [Medline: 29342478]
18. Lesho EP, Waterman PE, Chukwuma U, McAuliffe K, Neumann C, Julius MD, et al. The antimicrobial resistance monitoring and research (ARMoR) program: the US Department of Defense response to escalating antimicrobial resistance. Clin Infect Dis 2014 Aug 1;59(3):390-397. [doi: 10.1093/cid/ciu319] [Medline: 24795331]
19. Simões AS, Maia MR, Gregório J, Couto I, Asfeldt AM, Simonsen GS, et al. Participatory implementation of an antibiotic stewardship programme supported by an innovative surveillance and clinical decision-support system. J Hosp Infect 2018 Nov;100(3):257-264. [doi: 10.1016/j.jhin.2018.07.034] [Medline: 30071264]
20. Molloy JC. The Open Knowledge Foundation: open data means better science. PLoS Biol 2011 Dec;9(12):e1001195 [FREE Full text] [doi: 10.1371/journal.pbio.1001195] [Medline: 22162946]
21. RStudio. URL: https://www.rstudio.com/ [accessed 2018-11-16] [WebCite Cache ID 73ybbedpd]

XSL•FO
RenderX

22. R Core Team. R: A Language and Environment for Statistical Computing. URL: https://www.R-project.org/ [accessed 2018-11-16] [WebCite Cache ID 73ydeuUVX]
23. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. The Comprehensive R Archive Network. shiny: Web Application Framework for R URL: https://CRAN.R-project.org/package=shiny [accessed 2018-11-16] [WebCite Cache ID 73ydh83pW]
24. shinyapps.io. URL: https://www.shinyapps.io/ [accessed 2018-11-16] [WebCite Cache ID 73ybe2prd]
25. Stallman R. GNU General Public License, version 2. 1991. URL: https://www.gnu.org/licenses/gpl-2.0.html [accessed 2018-11-16] [WebCite Cache ID 73ydum67R]
26. Wickham H. Tidy data. J Stat Soft 2014;59(10):23. [doi: 10.18637/jss.v059.i10]
27. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse'. 2017. URL: https://www.tidyverse.org/ [accessed 2019-02-28] [WebCite Cache ID 76WntlTmp]
28. Grolemund G, Wickham H. Dates and times made easy with lubridate. J Stat Soft 2011;40(3):25. [doi: 10.18637/jss.v040.i03]
29. Zeileis A, Grothendieck G. zoo: S3 infrastructure for regular and irregular time series. J Stat Soft 2005;14(6):27. [doi: 10.18637/jss.v014.i06]
30. Berends MS, Luz CF, Glasner C, Friedrich AW, Sinha B. GitLab. AMR: An R package to simplify the analysis and prediction of Antimicrobial Resistance (AMR) and work with antibiotic properties by using evidence-based methods URL: https://gitlab.com/msberends/AMR [accessed 2019-05-09] [WebCite Cache ID 78FB9eqxz]
31. WHO Collaborating Centre for Drug Statistics Methodology. Oslo, Norway; 2018. Guidelines for ATC classification and DDD assignment URL: https://www.whocc.no/filearchive/publications/guidelines.pdf [accessed 2018-11-16] [WebCite Cache ID 73yeFdXyM]
32. Kassambara A, Kosinski M. The Comprehensive R Archive Network. 2018. survminer: Drawing Survival Curves using 'ggplot2' URL: https://cran.r-project.org/web/packages/survminer/index.html [accessed 2018-11-16] [WebCite Cache ID 73yeLlWLx]
33. Anhoej J. The Comprehensive R Archive Network. 2018. qicharts2: Quality Improvement Charts URL: https://CRAN.R-project.org/package=qicharts2 [accessed 2019-03-11] [WebCite Cache ID 76nVgPWG5]
34. Anhøj J. Diagnostic value of run chart analysis: using likelihood ratios to compare run chart rules on simulated data series. PLoS One 2015;10(3):e0121349 [FREE Full text] [doi: 10.1371/journal.pone.0121349] [Medline: 25799549]
35. RadaR shinyapps.io. URL: https://ceefluz.shinyapps.io/radar/ [accessed 2018-11-16] [WebCite Cache ID 73ybyEBHu]
36. RadaR GitHub. URL: https://github.com/ceefluz/radar [accessed 2018-11-16] [WebCite Cache ID 73yc2PgeW]
37. RadaR CSS GitHub. URL: https://github.com/ceefluz/radar/blob/master/www/radar_style.css [accessed 2018-11-16] [WebCite Cache ID 73ycOiZA5]
38. Pettit N, Han Z, Choksi A, Bhagat P, Pisano J. Using the Epic® Antimicrobial Stewardship (ASP) module to optimize antimicrobial stewardship interventions. Open Forum Infect Dis 2017;4(suppl_1):S263-S264. [doi: 10.1093/ofid/ofx163.580]
39. Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. J Am Med Inform Assoc 2018 Oct 1;25(10):1419-1428 [FREE Full text] [doi: 10.1093/jamia/ocy068] [Medline: 29893864]
40. Pirracchio R, Cohen MJ, Malenica I, Cohen J, Chambaz A, Cannesson M, ACTERREA Research Group. Big data and targeted machine learning in action to assist medical decision in the ICU. Anaesth Crit Care Pain Med 2018 Oct 16:-. [doi: 10.1016/j.accpm.2018.09.008] [Medline: 30339893]
41. Hahn E, Blazes D, Lewis S. Understanding How the. Health Secur 2016;14(1):13-18. [doi: 10.1089/hs.2015.0035] [Medline: 26889576]
42. Hong N, Prodduturi N, Wang C, Jiang G. Shiny FHIR: an integrated framework leveraging Shiny R and HL7 FHIR to empower standards-based clinical data applications. Stud Health Technol Inform 2017;245:868-872 [FREE Full text] [Medline: 29295223]

Abbreviations

AMR: antimicrobial resistance
AMS: antimicrobial stewardship
CSS: cascading style sheets
DDD: daily defined doses
DOT: days of therapy
EDs: emergency departments
EHR: electronic health record
ESGAP: European Society of Clinical Microbiology and Infectious Diseases Study Group for Antimicrobial Stewardship
LOS: length of stay
RadaR: Rapid analysis of diagnostic and antimicrobial patterns in R

XSL•FO
RenderX

Edited by G Eysenbach; submitted 17.11.18; peer-reviewed by A Staib, J van Gemert-Pijnen; comments to author 27.02.19; revised version received 12.03.19; accepted 24.03.19; published 24.05.19.

Please cite as:
Luz CF, Berends MS, Dik JWH, Lokate M, Pulcini C, Glasner C, Sinha B
Rapid Analysis of Diagnostic and Antimicrobial Patterns in R (RadaR): Interactive Open-Source Software App for Infection Management and Antimicrobial Stewardship
J Med Internet Res 2019;21(6):e12843
URL: https://www.jmir.org/2019/6/e12843/
doi:10.2196/12843
PMID:31199325

©Christian Friedemann Luz, Matthijs S Berends, Jan-Willem H Dik, Mariëtte Lokate, Céline Pulcini, Corinna Glasner, Bhanu Sinha. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 24.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Social Jetlag and Chronotypes in the Chinese Population: Analysis of Data Recorded by Wearable Devices

Zhongxing Zhang[1], PhD; Christian Cajochen[2,3], PhD; Ramin Khatami[1,4], MD

[1]Center for Sleep Medicine, Sleep Research and Epileptology, Clinic Barmelweid AG, Barmelweid, Switzerland

[2]Centre for Chronobiology, Psychiatric Hospital of the University of Basel, Basel, Switzerland

[3]Transfaculty Research Platform Molecular and Cognitive Neurosciences, University of Basel, Basel, Switzerland

[4]Department of Neurology, Inselspital, Bern University Hospital and University of Bern, Bern, Switzerland

Corresponding Author:
Zhongxing Zhang, PhD
Center for Sleep Medicine, Sleep Research and Epileptology
Clinic Barmelweid AG
Barmelweid,
Switzerland
Phone: 41 62 857 22 38
Fax: 41 62 857 22 25
Email: zhongxing.zhang@barmelweid.ch

Abstract

Background: Chronotype is the propensity for a person to sleep at a particular time during 24 hours. It is largely regulated by the circadian clock but constrained by work obligations to a specific sleep schedule. The discrepancy between biological and social time can be described as social jetlag (SJL), which is highly prevalent in modern society and associated with health problems. SJL and chronotypes have been widely studied in Western countries but have never been described in China.

Objective: We characterized the chronotypes and SJL in mainland China objectively by analyzing a database of Chinese sleep-wake pattern recorded by up-to-date wearable devices.

Methods: We analyzed 71,176 anonymous Chinese people who were continuously recorded by wearable devices for at least one week between April and July in 2017. Chronotypes were assessed (N=49,573) by the adjusted mid-point of sleep on free days (MSFsc). Early, intermediate, and late chronotypes were defined by arbitrary cut-offs of MSFsc <3 hours, between 3-5 hours, and >5 hours. In all subjects, SJL was calculated as the difference between mid-points of sleep on free days and work days. The correlations between SJL and age/body mass index/MSFsc were assessed by Pearson correlation. Random forest was used to characterize which factors (ie, age, body mass index, sex, nocturnal and daytime sleep durations, and exercise) mostly contribute to SJL and MSFsc.

Results: The mean total sleep duration of this Chinese sample is about 7 hours, with females sleeping on average 17 minutes longer than males. People taking longer naps sleep less during the night, but they have longer total 24-hour sleep durations. MSFsc follows a normal distribution, and the percentages of early, intermediate, and late chronotypes are approximately 26.76% (13,266/49,573), 58.59% (29,045/49,573), and 14.64% (7257/49,573). Adolescents are later types compared to adults. Age is the most important predictor of MSFsc suggested by our random forest model (relative feature importance: 0.772). No gender differences are found in chronotypes. We found that SJL follows a normal distribution and 17.07% (12,151/71,176) of Chinese have SJL longer than 1 hour. Nearly a third (22,442/71,176, 31.53%) of Chinese have SJL<0. The results showed that 53.72% (7127/13,266), 25.46% (7396/29,045), and 12.71% (922/7257) of the early, intermediate, and late chronotypes have SJL<0, respectively. SJL correlates with MSFsc (r=0.54, P<.001) but not with body mass index (r=0.004, P=.30). Random forest model suggests that age, nocturnal sleep, and daytime nap durations are the features contributing to SJL (their relative feature importance is 0.441, 0.349, and 0.204, respectively).

Conclusions: Our data suggest a higher proportion of early compared to late chronotypes in Chinese. Chinese have less SJL than the results reported in European populations, and more than half of the early chronotypes have negative SJL. In the Chinese population, SJL is not associated with body mass index. People of later chronotypes and long sleepers suffer more from SJL.

(J Med Internet Res 2019;21(6):e13482) doi:10.2196/13482

XSL•FO
RenderX

KEYWORDS

chronotypes; social jetlag; wearable devices; nap; cardiopulmonary coupling; sleep; big data

Introduction

Many human biological processes and behavioral functions exhibit 24-hour rhythms driven by endogenous circadian clocks, ranging from metabolism, hormone secretion, and body temperature to sleep-wake patterns and socialization. These endogenous daily variations among humans can be determined by "chronotype", which is controlled by circadian clocks but can be influenced by external environmental factors. Social jetlag (SJL) is the misalignment between biological and social time [1]. It is highly prevalent in modern society because of different factors such as increased workload and work stress in daily life, shift work, environmental changes [2] (eg, climate changes [3]), increasing exposure to artificial light sources, as well as excessive tablet/smartphone use [4]. SJL is reported to be associated with many health problems like obesity [5], increased daytime sleepiness and fatigue [6,7], bad mood and depression [8], and metabolic and cardiovascular disorders [9]. Reducing SJL is an important issue in public health.

One cause of SJL is insufficient sleep during work days, when people need to wake up earlier than their normal biological time and subsequently oversleep on free days to compensate for the accumulated sleep debt during the week [1,5]. Therefore, it is reasonable to assume that regular daytime naps to counteract accumulating sleep debt would alleviate SJL symptoms. Currently, most studies on SJL have been performed in Western countries [10,11]. It is difficult to investigate the relationship between daytime naps and SJL in these countries, since typically the general public rarely takes naps. Although "siesta" is historically common in Mediterranean countries such as Spain and Italy, Spaniards and Italians are abandoning this tradition and the majority of people no longer take naps [12]. In contrast, the noon nap is a Chinese tradition and is well protected in modern China despite the country's rapid industrialization and economic boom. Chinese people consider napping at noon as important for their health, and the noon nap is usually mandatory for kids at school. Chinese adults commonly take daily noon naps to refresh their energy for the activities in the afternoon and continue with this habit even after retirement [13]. Typically, Chinese people have a 2-hour break at noon that allows them to eat lunch and then take a nap.

To the best of our knowledge, the current states of SJL, chronotypes, and sleep in the Chinese population have never been quantified. Among several reasons for the lack of Chinese data is the absence of a Chinese translation and validation of questionnaires used for studying SJL and chronotypes, such as the Munich ChronoType Questionnaire (MCTQ) [14]. Some questionnaires evaluating chronotypes such as Morningness-Eveningness Questionnaire (MEQ) and its short form, the reduced MEQ, have been translated into Chinese [15,16] but lack sufficient validation. China is a huge country with 56 different ethnic groups, and for these groups, a countrywide validation has not yet been done. Certain studies [15,16] show inconsistent validation results of chronotypes in the Beijing region, which limits them from application in large population studies.

Wearable devices are becoming increasingly popular. Some have been validated for human sleep measurements in daily life situations, making it possible to assess sleep, chronotype, and SJL in the general population during daily life in an objective and longitudinal way. However, wearable devices measuring sleep based only on accelerometers overestimate sleep duration as they cannot really distinguish sleeping from lying quietly [17-21]. Recently, new wearables devices have been improved by adding the function of monitoring autonomic activities such as heart rate and pulse wave using photoplethysmography [22] or electrocardiogram sensors [23]. With the help of novel algorithms like machine learning [22,24] and cardiorespiratory sleep staging techniques (eg, cardiopulmonary coupling [CPC] [23,25] and heart rate variability analysis [26]), movement tracking (ie, accelerometer) and autonomic activities have improved the detection and determination of sleep and wakefulness [22,24,26]. Therefore, these new wearable devices are becoming more attractive in sleep and chronobiology research [23,27,28].

In this study, we characterize sleep, chronotypes, and SJL in mainland China by analyzing a large database of Chinese sleep-wake pattern recorded by up-to-date smartwatches. The device used in this study assesses the sleep-wake state by measuring movements and CPC with accelerometers and photoplethysmography. We hypothesize that the Chinese population has a smaller SJL compared to the data published in European countries, since napping at noon is very common in China. Given that insufficient sleep during the work day is one of the potential contributors to SJL [1,5], we reason that the Chinese "noon nap" culture is helpful for reducing daily sleep debt and counteracting SJL.

Methods

We analyzed 71,176 anonymous Chinese customers (male: 45,582 [64.04%], female: 9714 [13.65%], unknown: 15,880 [22.31%]) who use a wearable device from a major Chinese technology and telecommunications brand. For privacy reasons, we cannot reveal the name of the company. All customers were continuously recorded by their smartwatches for at least one week between April and July in 2017. Their nocturnal sleep durations were between 3 hours and 13 hours, that is, people who slept less than 3 hours or more than 13 hours were excluded. The age of these subjects was between 10 and 90 years old. We excluded customers with extremely high (>50 kg/m^2) or low (<10 kg/m^2) body mass indices (BMI). All subjects provided an electronic informed consent for their data to be used for research purposes when they first registered and initialized their devices. Hence, the data were collected naturally and safely during the customers' daily life. The wearable devices are able to distinguish sleep and wakefulness with high accuracy by combining movement tracking and CPC assessment. The accuracy of the wearable devices assessing sleep has been

XSL•FO
RenderX

validated by comparing with in-lab video-polysomnography [25,29]. This study was approved by the scientific and executive board of Clinic Barmelweid, Switzerland.

Mid-sleep point is the middle time between sleep onset and waking; for example, mid-sleep point is 4:00 am if sleep onset is at 00:00 and wake up time is 8:00 am in the morning. In each subject, we first calculated the mid-sleep points of all nights and then averaged the values on work days (from Sunday night to Thursday night) and on free days (Friday and Saturday nights), respectively. SJL is calculated as the averaged mid-sleep points on free days minus the averaged mid-sleep points on work days.

Changes of the adjusted mid-point of sleep on free days (ie, mid-sleep on free days corrected for sleep debt on work days) (MSFsc) were also characterized in 49,573 subjects (male: 30,651 [61.83%], female: 6394 [12.90%], unknown: 12,528 [25.27%]) whose data included at least 2 whole weeks with no missing data for age. In this way, we could test if the relationship between MSFsc and age as reported in a previous study [30] can be reproduced in the Chinese population. In addition, MSFsc measures human chronotype. Assessing chronotype is relevant for SJL because people of later chronotypes (ie, more evening type) are more likely to have higher sleep debt on work days and consequently a higher SJL [5]. The MSFsc was calculated according to the literature [30] as follows:

$$MSFsc = MSF - 0.5*[SD_f - (5*SD_w + 2*SD_f)/7]$$

where SD_f was the mean sleep duration of Friday and Saturday nights, and SD_w was the mean sleep duration from Sunday to Thursday nights. As these 49,573 subjects had recordings of whole weeks, MSFsc of each subject was the mean value of the MSFsc of recording weeks. For example, if a subject was recorded for 3 weeks, we calculated the MSFsc of each week and then the average of the MSFsc of these 3 weeks was the final MSFsc of this subject. Then we classified the chronotypes into early, intermediate, and late types with the arbitrary cut-offs (ie, early: <3:00 am intermediate: 3:00-5:00 am, late: >5:00 am) as suggested by Roenneber et al [31].

Three subgroups of subjects were defined according to the average durations of their daytime nap: (1) group with long naps, that is, average nap duration longer than 30 minutes (n=29,917), (2) group without daytime nap whose average nap duration was 0 (n=4223), and (3) group with short naps, that is, the rest of the subjects (n=37,036) taking daytime naps shorter than 30 minutes. A two-sample *t* test was used to check if SJL showed gender differences in all subjects and in the long naps and zero naps groups, respectively. The correlations between SJL and age/BMI/MSFsc were assessed by Pearson correlation in all subjects and in the long naps and zero naps groups, respectively. The short naps group was excluded from the aforementioned analysis because their mean nap duration was short and the nap effect may be minor in some subjects. Linear regression was used to quantify the relationship between SJL and daytime nap/nocturnal sleep durations. The data were expressed as mean ± standard deviation (SD), unless otherwise stated. The statistically significant level was $P<.001$, considering the large sample size.

We used random forest [32], a supervised machine learning approach that can be used for both classification and nonlinear regression analysis, to characterize which factors influence SJL and MSFsc. Random forest can handle highly nonlinear interactions between variables and does not require distributional assumptions (ie, normality) for the variables. The predictors included sex, age, BMI, the count of daily walking steps (an indicator of exercise), nocturnal sleep duration, and daytime nap duration. Although classical linear regression analysis was frequently used to estimate the relationships among variables, it was not suitable for our dataset due to the multicollinearity (ie, the covariates were highly correlated to each other). To eliminate the influences of missing data on the weights of the predictors, we fitted only the random forest model with complete dataset of all the predictors (n=32,020). The number of trees was set from 10 to 50 with a stepwise increment of 10, and the maximal depth of the tree was set from 2 to 5 with a stepwise increment of 1. We used a 5-folder cross-validation to select the optimal model, which gave the highest score (ie, the score was the coefficient of determination R-squared of the random forest regression). All the statistical analyses and random forest regression were done using Python.

Results

Sleep Duration

The distributions of age and BMI of our subjects are shown in Figure 1. The majority of our subjects were between 20 and 50 years old, with a BMI typically between 17 and 30 kg/m^2. On average, subjects (n=71,176) fell asleep at 00:15 am (SD 70.5 minutes) and woke up at 7:00 am (SD 66.3 minutes) on work days, while they fell asleep at 00:28 am (SD 82 minutes) and woke up at 7:22 am (SD 81.3 minutes) on free days. Thus, their mean nocturnal sleep durations were 387 minutes (SD 49.5) on work days and 396.6 minutes (SD 65.8) on free days. Their nocturnal sleep was 10 minutes longer on free days compared to work days (Welch's *t* test gives $P<.001$). The vast majority of our subjects took a daytime nap (66,953 out of 71,176 people, which was approximately 94.07% of our subjects). The distribution of the durations (>0 minutes) of the daytime naps is shown in Figure 2. Most of our subjects took naps shorter than 60 minutes. The total 24-hour sleep durations were 419 minutes (SD 47) (ie, nocturnal durations plus daytime nap durations), which were longer in females (433 minutes, SD 47) than males (416 minutes, SD 47) (two-sample *t* test, $P<.001$). We found a correlation between total sleep durations on work days and on free days in all subjects ($r=0.41$, $P<.001$), as well in the subgroups with zero naps ($r=0.43$, $P<.001$), with short naps ($r=0.42$, $P<.001$), and with long naps ($r=0.37$, $P<.001$), indicating that people who slept longer during work days also slept longer on weekends.

XSL•FO
RenderX

Figure 1. The distribution of age and body mass index (BMI) of our subjects.

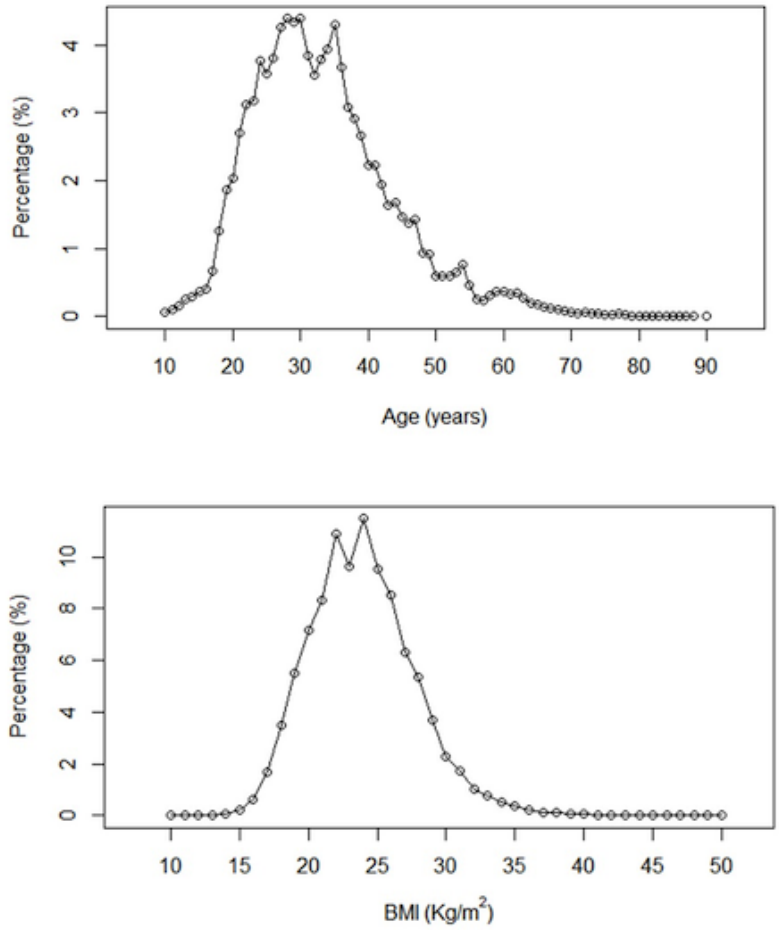
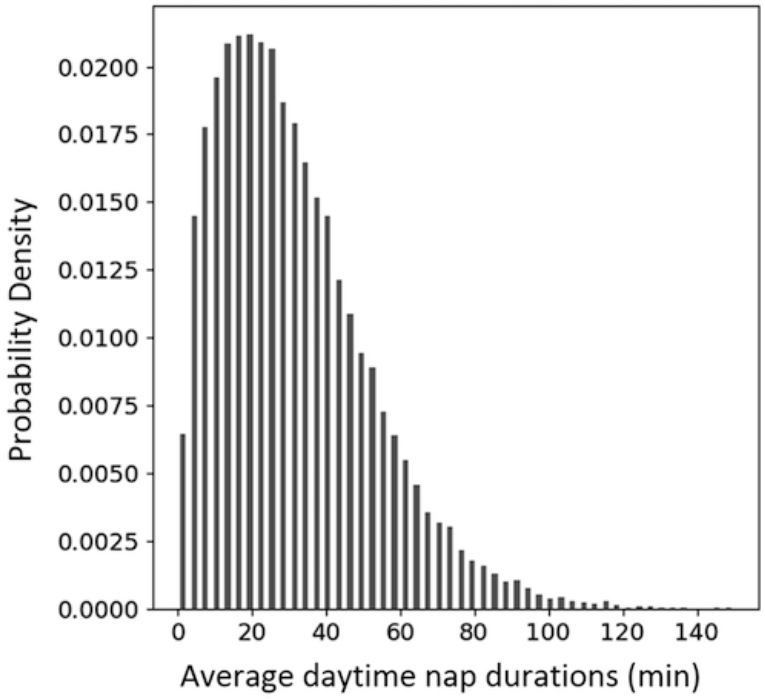


Figure 2. The distribution of the duration (>0 minutes) of daytime naps.



The nocturnal sleep duration negatively correlated with BMI (r=-0.216, P<.001) and with age (r=-0.14, P<.001). But no correlation was found between nap durations and BMI (r=0.003, P=.54). Further, nap durations did not correlate with age (r=0.05, P<.001), since the correlation coefficient was very small in spite of a small P value. It was thus not surprising to find negative

XSL•FO
RenderX

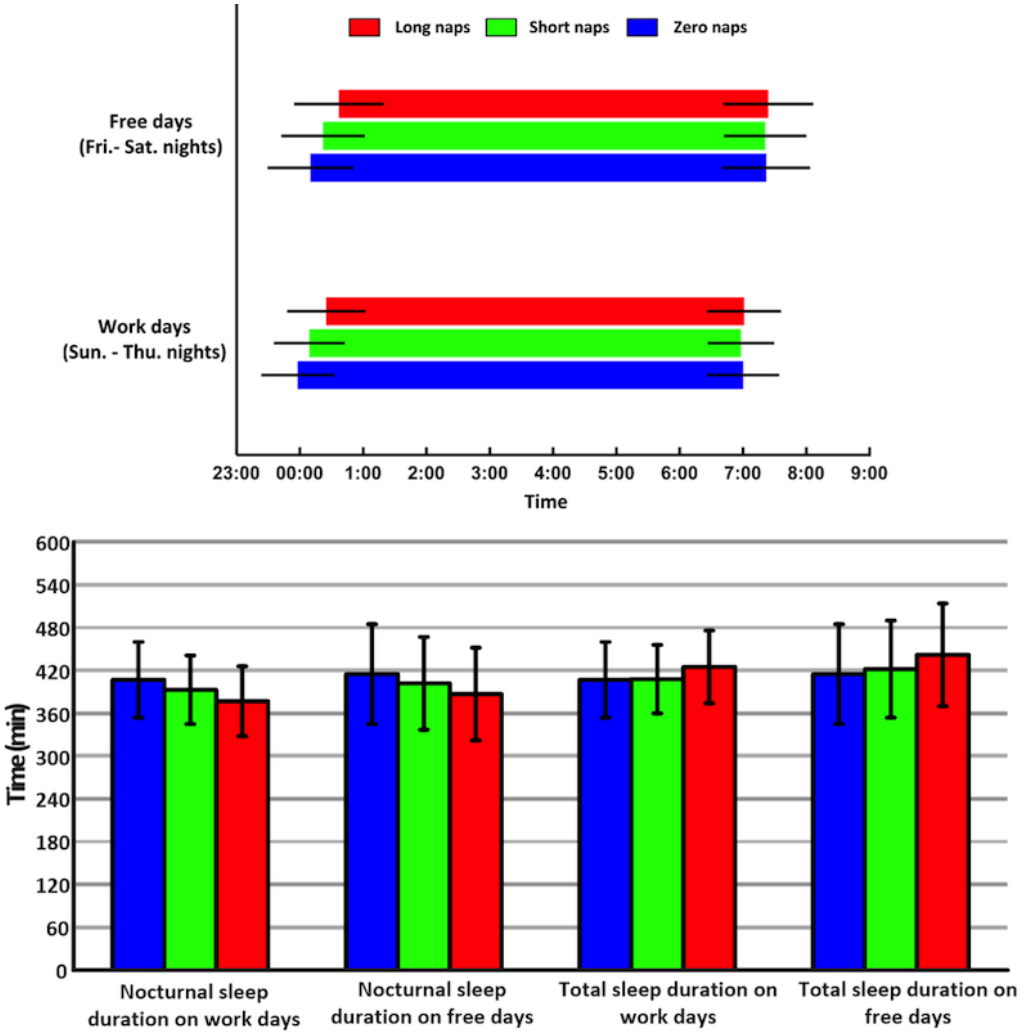
correlations between total sleep durations and BMI (r=-0.21, P<.001) as well as between total sleep durations and age (r=-0.12, P<.001), since nocturnal sleep duration was much longer than nap duration.

Figure 3 summarizes the sleep schedules of the three subgroups (ie, zero naps, short naps, and long naps). People taking longer daytime naps fell asleep later on both work and free days, but people in the three subgroups woke up at a similar time in the morning. The sleep schedules slightly shifted later on free days compared to work days in all the subgroups. After removing the durations of wake after sleep onset from sleep recordings, the average sleep durations of the three subgroups are illustrated in Figure 3. The group of zero naps had longer nocturnal sleep duration on both work and free days than the other two groups, but people taking longer daytime naps had significantly longer total sleep durations across the 24-hour day. This was confirmed by correlation analysis as nocturnal sleep duration was negatively correlated with nap duration (r=-0.22, P<.001).

Mid-Sleep on Free Days Corrected for Sleep Debt on Work Days

The distribution of MSFsc is shown in Figure 4. We found a near-normal distribution of chronotypes, with the most frequent MSFsc between 3:00 am and 4:00 am. The percentages of early (red color in Figure 4), intermediate (green color in Figure 4), and late (blue color in Figure 4) types were approximately 26.76% (13,266/49,573), 58.59% (29,045/49,573), and 14.64% (7257/49,573) in our subjects. The trajectory of age-dependent changes in MSFsc (Figure 5) showed a biphasic pattern with the peak value of 4:00 am at the age of 22 years. The MSFsc at 22 years was significantly later than the ones at ages younger than 16 years and older than 26 years (Welch's t test, P<.001). No gender differences were found (Welch's t test, P>.05) in all age groups except for the ones of 34 years (P<.001), 35 years (P<.001), 39 years (P<.001), and 40 years (P<.001).

Figure 3. The sleep schedules and sleep durations on work and free days in the 3 subgroups with zero naps, short naps (≤30 minutes), and long naps (>30 minutes). The error bars are standard deviation. The group of zero naps has longer nocturnal sleep duration on both work (P<.001) and free days (P<.001) than the other two groups, and the group of short naps has longer nocturnal sleep on both work (P<.001) and free days (P<.001) than the long naps group. However, people taking longer daytime naps have significantly longer total sleep durations across the 24-hour day on both work (P<.001) and free days (P<.001) than the other two groups.



XSL•FO
RenderX

Figure 4. The distribution of adjusted mid-point of sleep on free days (MSFsc) in our subjects (N=49,573). The red, green, and blue colors mark the early, intermediate, and late types using arbitrary cut-offs of <3 am, 3-5 am, and >5 am, respectively. h: hours.

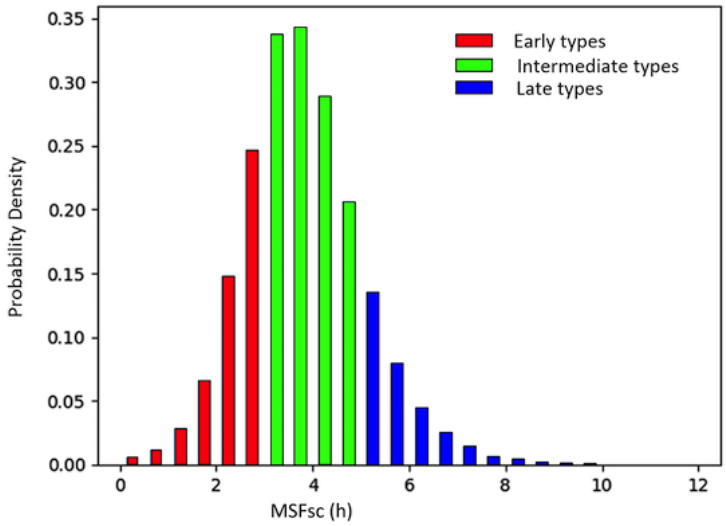
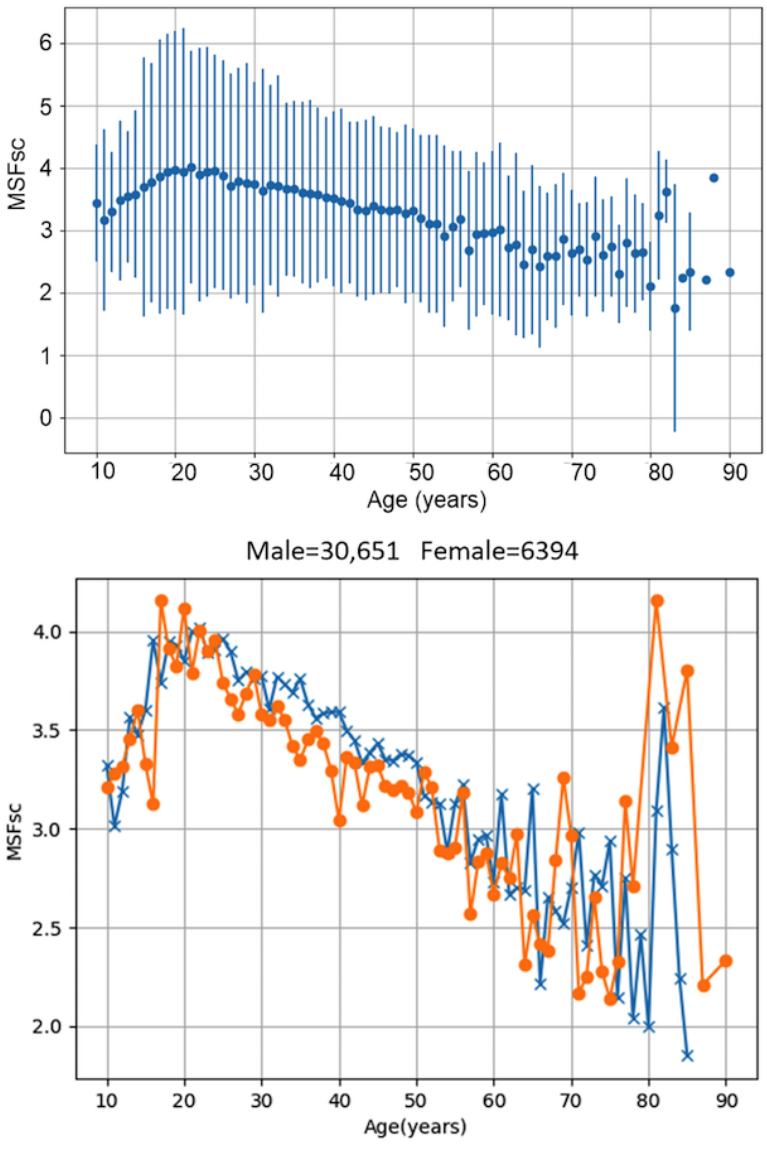


Figure 5. Age-dependent changes in adjusted mid-point of sleep on free days (MSFsc), and age-dependent changes of MSFsc in males and females (orange line with filled circles=female). The error bar is standard deviation.



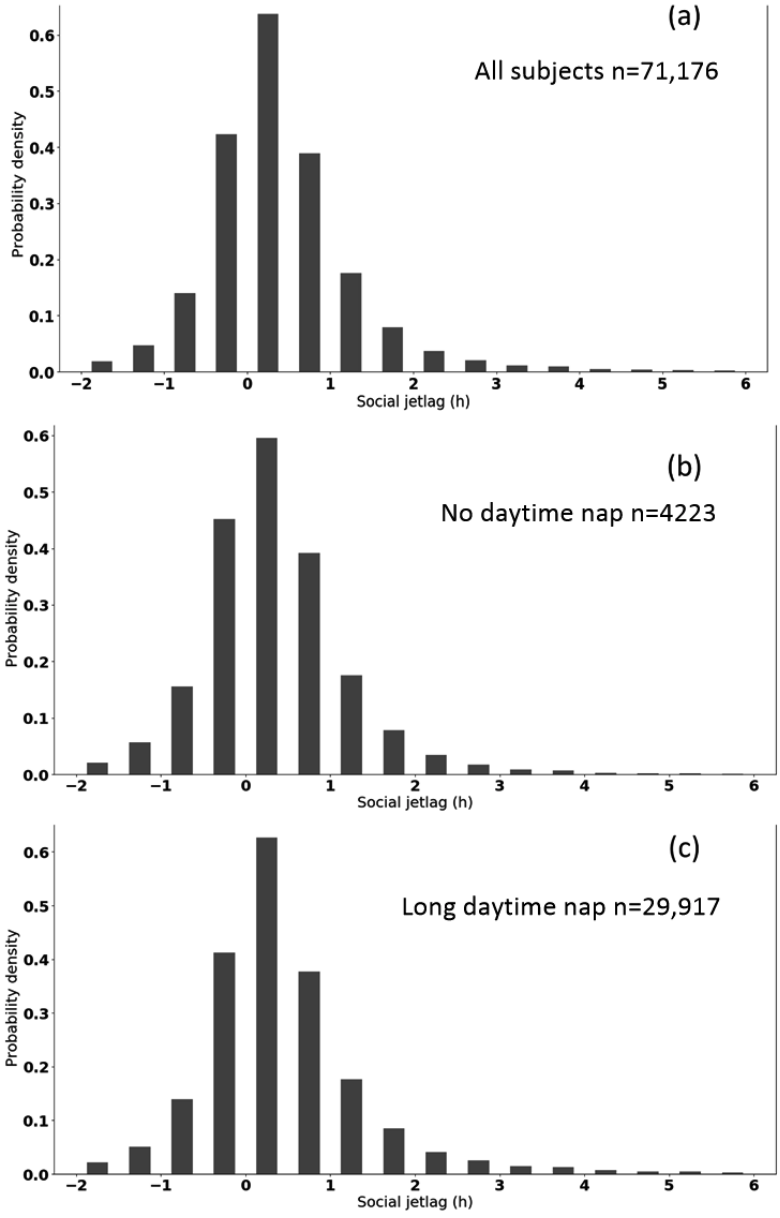
XSL•FO
RenderX

The optimal model of random forest regression was constructed with 50 trees with tree depth of 4. The relative feature importance for sex, age, BMI, exercise, nocturnal sleep duration, and daytime nap duration was 0.024, 0.772, 0.023, 0.036, 0.056, and 0.089, respectively, indicating that age was the predominant factor predicting MSFsc (ie, chronotype) in Chinese population. However, gender was not an important factor influencing MSFsc. The relative feature importance of nocturnal and daytime sleep durations suggested that they did not influence MSFsc in Chinese, as confirmed by correlation analyses. That is, nocturnal sleep (r=0.0048, P=.39) and nap durations (r=-0.008, P=.15) did not correlate to MSFsc.

Social Jetlag

Only 17.07% (12,151/71,176) of Chinese had SJL of more than 1 hour (see the distribution of SJL of our subjects in Figure 6). The distributions of SJL in the zero naps and long naps subgroups were similar to that of all the subjects. Figure 6 also shows that quite a number (22,442/71,176, 31.53%) of people had SJL<0. We further analyzed the number of negative SJL in different chronotypes (the chronotypes of 49,573 subjects were calculated as introduced above). We found 31.16% (15,447/49,573) of Chinese had negative SJL, that is, their mid-sleep points on work days were later than the ones on free days. Results showed that 46.14% (7127/15,447) of these subjects were early types, 47.88% (7396/15,447) intermediate types, and only 5.97% (922/15,447) late types. In other words, just over half (7127/13,266, 53.72%) of all the early types Chinese showed negative SJL. This proportion was 25.46% (7396/29,045) in the intermediate and 12.71% (922/7257) in the late types.

Figure 6. Histogram of social jetlag of (a) all subjects, (b) people with zero naps, and (c) people with long naps. h: hours.



XSL•FO
RenderX

Social jetlag did not correlate with BMI (all subjects without missing values of BMI: n=58,370, r=0.004, P=.30; subgroup with zero naps: n=3324, r=0.017, P=.34; subgroup with long naps: n=24,891, r=-0.007, P=.30). When further dividing our subjects into two subgroups (ie, BMI≥25 and BMI<25), no correlation was found between SJL and BMI in either group (BMI≥25: n=23,841, r=-0.009, P=.36; BMI<25: n=34,529, r=-0.002, P=.78). Taken together, these results indicate that SJL may not be associated with increased BMI. No gender difference was found in SJL in all subjects and the two subgroups of zero and long naps (male vs female in all subjects: 0.29 hours [SD 1.25] vs 0.31 hours [SD=1.02], P=.06; in the zero naps group: 0.26 hours [SD 1.19] vs 0.25 hours [SD 0.85], P=.86; in the long naps group: 0.3 hours [SD 1.42] vs 0.3 hours [SD 1.15], P=.77). Social jetlag significantly correlated to MSFsc in all subjects (r=0.54, P<.001) and in the two subgroups with zero daytime naps (r=0.56, P<.001) and with long daytime naps (r=0.58, P<.001). This result suggests that SJL was related to chronotypes, that is, people of later chronotypes may suffer more from SJL.

As MSFsc significantly correlated with SJL, we excluded it from the predictors in our random forest regression as we were interested in finding other predictors of SJL. Random forest regression showed that the best model was built with 40 trees with a depth of 2. The relative feature importance for sex, age, BMI, exercise, nocturnal sleep duration, and daytime nap duration was 0, 0.441, 0.006, 0, 0.349, and 0.204, respectively, suggesting that within the available predictors, age was the most important contributor to SJL, followed by the durations of nocturnal sleep and daytime nap. These results also suggest that sex and BMI were not relevant to SJL, confirming the classical correlation analysis and t test results mentioned above.

We next used classical statistics to explore the relationship between SJL and age/nocturnal sleep/daytime naps. Although the P values suggested highly significant results (P<.001), the relatively small correlation coefficients suggested that SJL did not correlate with age in all subjects (r=-0.03, P<.001). Linear regression analysis was used to quantify the relationship between SJL and daytime naps. The fitted slope of nap durations was 0.0066 and P<.001 (ie, if the daytime nap increased 1 minute, the SJL approximately increased 0.0066*60=0.4 minutes). When adding nocturnal sleep duration into the regression model, the fitted slope of daytime nap was not significant any more (P=.16) but nocturnal sleep duration became significant (the fitted slope was 0.0007 and P<.001). These results suggested a positive relationship between SJL and durations of daytime nap/nocturnal sleep, indicating that people who need more sleep seemed to have larger SJL.

Discussion

Principal Considerations

This is the first study characterizing sleep duration, chronotype (ie, mid-sleep points on free days), and social jetlag in a large Chinese population with objective data assessed by wearable devices. Our results reveal that the majority of Chinese people take daytime naps (66,953/71,176, 94.07%), and most importantly only a minor proportion (12,151/71,176, 17.07%) of the Chinese population have SJL longer than 1 hour, which is much smaller than the 69% reported in a European population [5]. Remarkably, in contrast to the results shown in previous studies that SJL is associated with obesity [5,33,34], we report that SJL does not associate with BMI in a Chinese population. Surprisingly, a high proportion of Chinese people, that is, 31.53% (22,442/71,176) of all the subjects or 31.16% (15,447/49,573) of the subjects whose chronotypes are available, have negative SJL. Among these people, the majorities are early (7127/15,447, 46.14%) and intermediate (7396/15,447, 47.88%) chronotypes. In fact, we found that more than half (7127/13,266, 53.72%) of the whole early types show negative SJL, indicating that negative SJL may be a common phenomenon in people of early chronotypes. In contrast to our original hypothesis that daytime naps reduce SJL, we found that people sleeping longer during daytime naps also have a larger SJL. We further showed that chronotype (ie, MSFsc), age, and total sleep duration (ie, the sum of nocturnal and daytime sleep durations) are the most important factors associating with SJL. People of later chronotypes and long sleepers have a larger SJL. In addition, we report a higher proportion of early chronotypes (13,266/49,573, 26.76%) than late chronotypes (7257/49,573, 14.64%) and show the age-MSFsc relationship in Chinese population as previously reported in the Western countries [30,35,36], confirming that MSFsc becomes later during adolescence. However, its peak is earlier than the one reported in the German-speaking countries in 2004 [30] (ie, 4:00 am vs 5:00 am), indicating that nowadays the shift of MSFsc towards later chronotypes during adolescence is smaller compared to 10 years ago. We did not find gender differences in MSFsc.

Sleep Duration

The average nocturnal sleep duration is less than 7 hours in our subjects, which is shorter than that reported in other countries like the United States [,] and European countries [37]. In fact, at least 7 hours of sleep is recommended by the US Centers for Disease Control and Prevention to promote optimal health [38]. However, the majority of the Chinese population taking daytime naps results in an average total sleep duration of 7 hours (ie, 419 minutes [SD 47 minutes]). Our results also show gender differences in total sleep duration, that is, females sleep on average 17 minutes longer than males, which is consistent with the results of previous studies [39,40].

Early Chronotypes in the Chinese Population

This is the first study that chronotyped the Chinese population with an objective assessment. Very few previous studies have evaluated the chronotypes of Chinese using subjective questionnaires, and their results have been inconsistent. For example, Li et al [15] and Carciofo et al [16] validated the MEQ questionnaire with a sample of 188 subjects from the Beijing district and a sample of 305 Beijing residents. They found inconsistent results of chronotypes classified with the Horne and Östberg cut-off criteria [41]. Carciofo et al reported that 48.9% of subjects were neutral type, 44.3% were definitely or moderately morning types, and 6.9% definitely or moderately evening types [16]. Li et al showed that 66% of subjects were definitely or moderately morning types, 31% were neutral type, and 3% definitely or moderately evening types [15]. However,

XSL•FO
RenderX

both studies indicated a prevalence of early type compared to late chronotypes in China. In our results, the objective MSFsc in the Chinese population follows a near normal distribution (Figure 4) similar to the ones shown in a recent study done in the US with a questionnaire [35]. With the same arbitrary cut-offs used in the German-speaking populations [31] and in the US population [35], we also found a higher proportion of early than late types. Therefore, our results of objective assessment of chronotypes agree with the findings of previous studies in Chinese using subjective questionnaires.

Age Differences in Chronotypes

Our results (Figure 5) show that Chinese adolescents tend to be later chronotypes than other age groups consistent with previous results reported in other countries such as German-speaking countries [30,36] and the United States [35]. The big advantage of our study is that the MSFsc was calculated from objective measurement of sleep schedules by wearable devices, whereas in previous studies only subjective MSFsc derived from questionnaires was available. Both objective and subjective assessments corroborate a later chronotype in adolescents. The objectively assessed peak of age related to the MSFsc curve is 4:00 am at the age of 22 years in Chinese and is 5:00 am at age 20 years in the German-speaking population assessed by the MCTQ in 2004 [30]. Thus, our results suggest that although currently the MSFsc still shifts towards later chronotypes in adolescents, the degree of the shift may be decreasing compared to 15 years ago or may be smaller when objectively assessed with wearable devices. This hypothesis is supported by data from the United States [35], as the authors compare the MSFsc of their survey years (2003-2014) and find a trend towards earlier chronotypes in the later survey years (2011-2014) compared to the early ones.

The peak value of MSFsc at age 22 years is not significantly different from the ones at the ages 16-26 years. This range of age is wider than the one shown in the German-speaking population (ie, 19-23 years) [30]. The 24-hour sleep-wake cycle is influenced by genetic, molecular, lifestyle, societal (eg, cultures, working hours, occupation), environmental (eg, pollution, climate), geographical, and other factors. Thus, these factors may also influence the calculation of the MSFsc (the chronotypes) [42]. As Roenneberg et al suggested that the peak of age-MSFsc curve is a marker for the end of adolescence [30], the genetic factors may account for the aforementioned differences among different countries. Other factors like societal, environmental, and geographical factors could also influence the age-MSFsc relationships. For example, China covers a larger range of latitude (ie, from the northern tropic to high latitudes) and longitude than the German-speaking countries (eg, Germany and Switzerland are at high latitudes). Higher latitudes are associated with higher eveningness [10]. Therefore, when we average the MSFsc of Chinese people across the whole country, it is plausible that the Chinese have a smaller value but wider age ranges of the latest MSFsc compared to the German-speaking population.

Gender Differences in Chronotypes

In contrast to the results of previous studies, our objective data do not confirm gender differences in the Chinese population. In fact, to the best of our knowledge, the gender differences in chronotypes are reported inconsistently in the literature. Large sample and meta-analysis studies have found a slight tendency towards early chronotypes in females compared to males [30,35,36,43,44], but some studies report controversial findings [45,46] or no gender differences [43]. For example, a study [45] comparing the Morningness-Eveningness Stability Scale among Iranian, Spanish, and German populations found that the morning affect is higher in Spaniards and German females compared to Spaniards and German males. Iranian females, however, reported lower morning affect scores compared to Iranian males, indicating that gender differences in chronotypes are controversial between Iranian and Spaniard/German populations. In contrast to the results of most studies done in Europe showing gender differences [36,44], Duarte et al reported no significant gender differences in age groups between 30 and 44 years in Brazil using the MEQ questionnaire [47]. In Finland, conflicting results were also reported [46]. Evening type is more common among women than men in the Finnish population.

Negative Social Jetlag

MSFsc is positively correlated to SJL, corroborating the notion that people of later types may have higher SJL [5] because currently human society in industrialized countries favors early over later working hours. However, our results also report negative SJL in quite a number of Chinese, especially in early types. Negative SJL has been reported in previous studies [5], and Roenneberg et al suggested that this may be because people of early types sleep later than their circadian sleep window during work days, thus on free days they go to sleep earlier [5]. Other possible explanations of the negative SJL include shift workers who work on weekends, or accumulated sleep pressures forcing some people to go to sleep earlier on free days (ie, sleeping in is not the only way to compensate for insufficient sleep; some people also prefer to sleep earlier on free days). Nevertheless, our results report for the first time a relative high prevalence of negative SJL in a larger population (ie, Chinese) with a higher proportion of early chronotypes than late chronotypes. To the best of our knowledge, the health consequences of negative SJL have been rarely investigated in previous studies, probably due to the relative low proportion of early chronotypes/negative SJL in the studied populations (eg, the absolute values of SJL were calculated as very few subjects reported negative SJL in Roenneberg et al [5]). Only one recent study done in Japan shows that about 22.3% of Japanese children and adolescents report negative SJL and the students with negative SJL less than 1 hour have poorer academic performance [48]. Therefore, we suggest that more data are needed to explore the potential influences of negative SJL in human health, especially for early chronotypes whose sleep windows may be delayed during the work days. These data will help us better understand how the misalignments of biological and social time influence human health. Previous studies emphasized the influences of SJL on late chronotypes but ignored its influences on early chronotypes. For example, late chronotypes suffer from earlier waking in the morning, so later school/work start times have been recommended in some countries. But early chronotypes may equally suffer from later sleep onset during the night and their circadian clock prefers an early school/work

XSL•FO
RenderX

start. So, reducing the negative SJL should be equally important for public health, especially for early chronotypes.

Age and Social Jetlag

Age is suggested to be associated with SJL by our random forest model, but it is not confirmed by classical correlation analysis as the correlation coefficient is too small (r=-0.03) in spite of small P value (P<.001). This result could be explained by the nonlinear relationship between age and SJL, similar to the one between age and MSFsc. As shown in our results (Figure 5), children are early chronotypes and become progressively later in adolescence. Chronotypes are positively correlated to SJL, so SJL is increasing with maturation in childhood. In adulthood, MSFsc is decreasing with increased age (ie, chronotypes become earlier again); thus, SJL is decreasing with age. This nonlinear pattern cannot be discovered by linear correlation analysis, but it can be recognized by random forest, a decision tree–based algorithm suitable to discovery of nonlinear relationships between predictors and dependent variable.

Body Mass Index and Social Jetlag

In contrast to the results of previous studies [5,33,34], we could not find the association between SJL and BMI in the Chinese population. In fact, we are not the only study reporting no association between SJL and BMI. Similar results have been recently reported in Russia [49], Norway [50], and Czech Republic [51]. Thus, societal or geographical (eg, Russia and Norway are of higher latitudes compared to Germany) factors could account for the inconsistent results. Different data analysis methods may also influence the results. For example, both our analyses with machine learning (ie, random forest) and correlation analysis failed to reveal an association between SJL and BMI in our data. However, if we used ordinary least squares regression with BMI, gender, age, exercise, and sleep duration as predictors, we could find BMI as a positive predictor (coefficient is 0.0033, and P=.017) for SJL. However, this regression analysis is problematic due to the multicollinearity of covariates. In addition, energy expenditure and energy intake have not been considered in any of these studies including this one. Although we have taken into account the count of daily walking steps measured by wearable devices, which is an indicator of exercise, we have no indicator of energy intake. Therefore, the link between SJL and obesity needs further research.

Daytime Naps and Social Jetlag

The unexpected positive rather than negative relationship between SJL and daytime naps is best explained by higher sleep need. First, our data show that the total sleep durations on work days and free days are correlated (r=0.41, P<.001), indicating that those people sleeping longer are long sleepers while the people without daytime naps have shorter total sleep durations and are likely to be short sleepers. Second, our result of an average of 7 hours total sleep duration already indicates that in general Chinese people suffer from insufficient sleep. People taking longer daytime naps sleep less during the night but have longer total sleep time (Figure 3), suggesting that they show a disposition to long sleep need (ie, long sleepers) but suffer from insufficient nocturnal sleep. Therefore, they sleep longer on free days, resulting in larger SJL (ie, longer daytime nap is associated with larger SJL). These results also indicate that long sleepers have larger SJL.

The positive association between daytime naps and SJL should not be considered to reflect a causal relationship and misinterpreted in that a longer daytime nap leads to larger SJL. In contrast, we think that daytime naps are still helpful to reduce SJL. However, they are not long enough to totally compensate for insufficient nocturnal sleep because the majority of Chinese people take daytime naps and have less than 1-hour SJL compared to European populations (ie, assuming that if Chinese people did not take daytime naps, they are likely to have higher sleep debts and thus larger SJL).

We also recognize that long daytime naps may also be associated with other health problems such as diabetes [52] or cardiovascular diseases [53]. Previous studies suggested that long daytime naps (ie, longer than 1 hour) can significantly increase the odds ratio of diabetes mellitus [52,54] and are associated with a higher risk of cardiovascular disease [53], but short naps (less than 30 minutes) are suitable to promote health. The influences/impacts of daytime napping duration on human health are less conclusive and definitely deserve more study. For example, long daytime naps may indicate sleep deprivation or disturbed nocturnal sleep (eg, people with sleep disorders such as sleep apnea, parasomnia, or insomnia). So, the increased risk of cardiovascular disease or diabetes in those people may be due to their poor nocturnal sleep or sleep disorders, rather than long daytime naps.

Limitations

There are several limitations to our study. First, we are well aware that the cut-offs of MSFsc classifying chronotypes are sensitive to the study population [55] and other factors like latitude and longitude [10]; thus, they should be interpreted with caution. Our results that nearly half of negative SJL are intermediate chronotypes may indicate that the proportion of early chronotypes in this study is underestimated using the cut-offs.

Second, we have no data on the subjective estimations of sleep. To correlate the objective measures and subjective evaluations (eg, sleep schedule, daytime sleepiness) may be an interesting topic in the future. Third, in China the majority of employees can work only under the standard working hour system, which limits them to work 8 hours per day and 40 hours per week. Therefore, most Chinese people work from Monday to Friday. But working on Saturday or Sunday (not on both days as it is forbidden to work for the whole week according to Chinese law) with overtime payment is allowed in China. In this case, the definition of free days in our study needs to be adjusted in a small proportion of subjects. We suggest that future studies should take this into account. For example, the questionnaires used to assess chronotypes and SJL should include working hours/days per week of the subjects. The objective recordings with wearable devices or mobile phones should be able to classify work days and free days. In addition, we are aware that seasonal effects can influence chronotypes [36] and they should be controlled as suggested in previous studies [56]. Therefore, we recorded our subjects from April to July, when the

XSL•FO
RenderX

temperature is warm and the duration of daytime sunshine is increasing. It would be interesting in the future to compare our results with the data measured by wearable devices from winter when the temperature is cold and the night is longer. These objective data collected from large populations will help us better understand how daylight and temperature influence human circadian rhythm and health. For example, they can contribute to the ongoing debate of whether we should stop shifting to daylight saving time in summer in Europe and North America and instead use permanent summer time or winter time.

Finally, we have no information of the timing but only the durations of daytime naps in the database. It is reasonable to speculate that the majority of our subjects may regularly take naps after lunch between 12:00 to 14:00 (typical midday break time in China) because noon napping is an intrinsic part of Chinese culture. But we acknowledge that this speculation needs to be tested in future studies. It is an interesting topic to investigate whether the timing or regularity of naps may influence nocturnal sleep or the calculation of MSFsc and SJL.

Conclusions

Our study characterizes for the first time the chronotypes and social jetlag in a large Chinese population measured using wearable devices. We found a higher proportion of early compared to late chronotypes in Chinese. Chinese had less SJL than the results reported in European populations probably because of their noon napping culture. We also found that people of later chronotypes and long sleepers suffered more from SJL, but larger SJL was not associated with higher BMI. Surprisingly, more than half of the early chronotypes had negative SJL. We suggest that future studies are needed to further investigate the relationships between negative SJL and human health. Our study demonstrates that these days the modern wearable technologies tracking sleep-wake patterns are becoming powerful tools in the field of chronobiology.

Acknowledgments

This work was supported by Clinic Barmelweid Scientific Foundation.

Conflicts of Interest

None declared.

References

1. Wittmann M, Dinich J, Merrow M, Roenneberg T. Social jetlag: misalignment of biological and social time. Chronobiol Int 2006;23(1-2):497-509. [doi: 10.1080/07420520500545979] [Medline: 16687322]
2. Amaral FG, Castrucci AM, Cipolla-Neto J, Poletini MO, Mendez N, Richter HG, et al. Environmental control of biological rhythms: effects on development, fertility and metabolism. J Neuroendocrinol 2014 Sep;26(9):603-612. [doi: 10.1111/jne.12144] [Medline: 24617798]
3. Obradovich N, Migliorini R, Mednick SC, Fowler JH. Nighttime temperature and human sleep loss in a changing climate. Sci Adv 2017 May;3(5):e1601555 [FREE Full text] [doi: 10.1126/sciadv.1601555] [Medline: 28560320]
4. Swaminathan K, Klerman EB, Phillips AJK. Are Individual Differences in Sleep and Circadian Timing Amplified by Use of Artificial Light Sources? J Biol Rhythms 2017 Dec;32(2):165-176 [FREE Full text] [doi: 10.1177/0748730417699310] [Medline: 28367676]
5. Roenneberg T, Allebrandt KV, Merrow M, Vetter C. Social jetlag and obesity. Curr Biol 2012 May 22;22(10):939-943 [FREE Full text] [doi: 10.1016/j.cub.2012.03.038] [Medline: 22578422]
6. Komada Y, Breugelmans R, Drake CL, Nakajima S, Tamura N, Tanaka H, et al. Social jetlag affects subjective daytime sleepiness in school-aged children and adolescents: A study using the Japanese version of the Pediatric Daytime Sleepiness Scale (PDSS-J). Chronobiol Int 2016;33(10):1311-1319. [doi: 10.1080/07420528.2016.1213739] [Medline: 27715324]
7. Martin JS, Hébert M, Ledoux E, Gaudreault M, Laberge L. Relationship of chronotype to sleep, light exposure, and work-related fatigue in student workers. Chronobiol Int 2012 Apr;29(3):295-304. [doi: 10.3109/07420528.2011.653656] [Medline: 22390242]
8. Levandovski R, Dantas G, Fernandes LC, Caumo W, Torres I, Roenneberg T, et al. Depression scores associate with chronotype and social jetlag in a rural population. Chronobiol Int 2011 Nov;28(9):771-778. [doi: 10.3109/07420528.2011.602445] [Medline: 21895489]
9. Wong PM, Hasler BP, Kamarck TW, Muldoon MF, Manuck SB. Social Jetlag, Chronotype, and Cardiometabolic Risk. J Clin Endocrinol Metab 2015 Dec;100(12):4612-4620 [FREE Full text] [doi: 10.1210/jc.2015-2923] [Medline: 26580236]
10. Randler C, Rahafar A. Latitude affects Morningness-Eveningness: evidence for the environment hypothesis based on a systematic review. Sci Rep 2017 Dec 03;7:39976 [FREE Full text] [doi: 10.1038/srep39976] [Medline: 28045131]
11. Adan A, Archer SN, Hidalgo MP, Di Milia L, Natale V, Randler C. Circadian typology: a comprehensive review. Chronobiol Int 2012 Nov;29(9):1153-1175. [doi: 10.3109/07420528.2012.719971] [Medline: 23004349]
12. Ohayon MM, Priest RG, Zulley J, Smirne S, Paiva T. Prevalence of narcolepsy symptomatology and diagnosis in the European general population. Neurology 2002 Jun 25;58(12):1826-1833. [Medline: 12084885]

XSL•FO
RenderX

13. Li J, Cacchione PZ, Hodgson N, Riegel B, Keenan BT, Scharf MT, et al. Afternoon Napping and Cognition in Chinese Older Adults: Findings from the China Health and Retirement Longitudinal Study Baseline Assessment. J Am Geriatr Soc 2017 Feb;65(2):373-380. [doi: 10.1111/jgs.14368] [Medline: 27995615]
14. Roenneberg T, Wirz-Justice A, Merrow M. Life between clocks: daily temporal patterns of human chronotypes. J Biol Rhythms 2003 Feb;18(1):80-90. [doi: 10.1177/0748730402239679] [Medline: 12568247]
15. Li S, Li Q, Wang X, Liu L, Liu Y, Zhang L, et al. Preliminary test for the Chinese version of the Morningness-Eveningness Questionnaire. Sleep and Biological Rhythms 2011;9(1):19-23. [doi: 10.1111/j.1479-8425.2010.00480.x]
16. Carciofo R, Du F, Song N, Qi Y, Zhang K. Age-related chronotype differences in Chinese, and reliability assessment of a reduced version of the Chinese Morningness-Eveningness Questionnaire. Sleep and Biological Rhythms 2012;10(4):-. [doi: 10.1111/j.1479-8425.2012.00577.x]
17. Kushida C, Chang A, Gadkary C, Guilleminault C, Carrillo O, Dement WC. Comparison of actigraphic, polysomnographic, and subjective assessment of sleep parameters in sleep-disordered patients. Sleep Med 2001 Sep;2(5):389-396. [Medline: 14592388]
18. Pollak C, Tryon W, Nagaraja H, Dzwonczyk R. How accurately does wrist actigraphy identify the states of sleep and wakefulness? Sleep 2001 Dec 15;24(8):957-965. [Medline: 11766166]
19. Paquet J, Kawinska A, Carrier J. Wake detection capacity of actigraphy during sleep. Sleep 2007 Oct;30(10):1362-1369 [FREE Full text] [Medline: 17969470]
20. de Zambotti M, Baker FC, Colrain IM. Validation of Sleep-Tracking Technology Compared with Polysomnography in Adolescents. Sleep 2015 Sep 01;38(9):1461-1468 [FREE Full text] [doi: 10.5665/sleep.4990] [Medline: 26158896]
21. Poyares D, Hirotsu C, Tufik S. Fitness Tracker to Assess Sleep: Beyond the Market. Sleep 2015 Sep 01;38(9):1351-1352 [FREE Full text] [doi: 10.5665/sleep.4966] [Medline: 26237767]
22. Fonseca P, Weysen T, Goelema M, Møst EIS, Radha M, Lunsingh Scheurleer C, et al. Validation of Photoplethysmography-Based Sleep Staging Compared With Polysomnography in Healthy Middle-Aged Adults. Sleep 2017 Dec 01;40(7):-. [doi: 10.1093/sleep/zsx097] [Medline: 28838130]
23. Thomas R, Wood C, Bianchi MT. Cardiopulmonary coupling spectrogram as an ambulatory clinical biomarker of sleep stability and quality in health, sleep apnea and insomnia. Sleep 2017 Dec 09;41(2):- [FREE Full text] [doi: 10.1093/sleep/zsx196] [Medline: 29237080]
24. Reimer U, Emmenegger S, Maier E, Zhang Z. Recognizing Sleep Stages with Wearable Sensors in Everyday Settings. 2017 Presented at: 3rd International Conference on Information and Communication Technologies for Ageing Well and e-Health - Volume 1: ICT4AWE; April 28-29, 2017; Porto, Portugal. [doi: 10.5220/0006346001720179]
25. Thomas R, Mietus J, Peng C, Goldberger AL. An electrocardiogram-based technique to assess cardiopulmonary coupling during sleep. Sleep 2005 Sep;28(9):1151-1161. [Medline: 16268385]
26. Domingues A, Paiva T, Sanches JM. Hypnogram and sleep parameter computation from activity and cardiovascular data. IEEE Trans Biomed Eng 2014 Jun;61(6):1711-1719. [doi: 10.1109/TBME.2014.2301462] [Medline: 24845281]
27. Fagherazzi G, El Fatouhi D, Bellicha A, El Gareh A, Affret A, Dow C, et al. An International Study on the Determinants of Poor Sleep Amongst 15,000 Users of Connected Devices. J Med Internet Res 2017 Dec 23;19(10):e363 [FREE Full text] [doi: 10.2196/jmir.7930] [Medline: 29061551]
28. Wen D, Zhang X, Liu X, Lei J. Evaluating the Consistency of Current Mainstream Wearable Devices in Health Monitoring: A Comparison Under Free-Living Conditions. J Med Internet Res 2017 Dec 07;19(3):e68 [FREE Full text] [doi: 10.2196/jmir.6874] [Medline: 28270382]
29. Zhang Z, Henzmann S, Hügli G, Qi M, Chen W, Lu C, et al. Validation of Wearable Sleep Monitoring Device Based on Cardiopulmonary Coupling and Accelerometer with Comparison to Polysomnography in Adults. : European Journal of Neurology; 2018 Presented at: 4th Congress of the European Academy of Neurology; June 16-19, 2018; Lisbon, Portugal p. 385 URL: https://tinyurl.com/yyq7aw6n
30. Roenneberg T, Kuehnle T, Pramstaller PP, Ricken J, Havel M, Guth A, et al. A marker for the end of adolescence. Curr Biol 2004 Dec 29;14(24):R1038-R1039 [FREE Full text] [doi: 10.1016/j.cub.2004.11.039] [Medline: 15620633]
31. Roenneberg T, Kuehnle T, Juda M, Kantermann T, Allebrandt K, Gordijn M, et al. Epidemiology of the human circadian clock. Sleep Med Rev 2007 Dec;11(6):429-438. [doi: 10.1016/j.smrv.2007.07.005] [Medline: 17936039]
32. Breiman L. Random Forests. Machine Learning 2001;45(1):5-32 [FREE Full text] [doi: 10.1023/A:1010933404324]
33. Parsons MJ, Moffitt TE, Gregory AM, Goldman-Mellor S, Nolan PM, Poulton R, et al. Social jetlag, obesity and metabolic disorder: investigation in a cohort study. Int J Obes (Lond) 2015 May;39(5):842-848 [FREE Full text] [doi: 10.1038/ijo.2014.201] [Medline: 25601363]
34. Mota MC, Silva CM, Balieiro LCT, Fahmy WM, Crispim CA. Social jetlag and metabolic control in non-communicable chronic diseases: a study addressing different obesity statuses. Sci Rep 2017 Dec 25;7(1):6358 [FREE Full text] [doi: 10.1038/s41598-017-06723-w] [Medline: 28743872]
35. Fischer D, Lombardi DA, Marucci-Wellman H, Roenneberg T. Chronotypes in the US - Influence of age and sex. PLoS One 2017;12(6):e0178782 [FREE Full text] [doi: 10.1371/journal.pone.0178782] [Medline: 28636610]

XSL•FO
RenderX

36. Randler C, Freyth-Weber K, Rahafar A, Florez Jurado A, Kriegs JO. Morningness-eveningness in a large sample of German adolescents and adults. Heliyon 2016 Nov;2(11):e00200 [FREE Full text] [doi: 10.1016/j.heliyon.2016.e00200] [Medline: 27957554]
37. Zomers M, Hulsegge G, van Oostrom SH, Proper K, Verschuren W, Picavet HJ. Characterizing Adult Sleep Behavior Over 20 Years-The Population-Based Doetinchem Cohort Study. Sleep 2017 Dec 01;40(7):- [FREE Full text] [doi: 10.1093/sleep/zsx085] [Medline: 28525637]
38. Liu Y, Wheaton A, Chapman D, Cunningham T, Lu H, Croft JB. Prevalence of Healthy Sleep Duration among Adults--United States, 2014. MMWR Morb Mortal Wkly Rep 2016 Feb 19;65(6):137-141 [FREE Full text] [doi: 10.15585/mmwr.mm6506a1] [Medline: 26890214]
39. Reyner L, Horne J, Reyner A. Gender- and age-related differences in sleep determined by home-recorded sleep logs and actimetry from 400 adults. Sleep 1995 Feb;18(2):127-134. [Medline: 7792492]
40. Goel N, Kim H, Lao RP. Gender differences in polysomnographic sleep in young healthy sleepers. Chronobiol Int 2005;22(5):905-915. [doi: 10.1080/07420520500263235] [Medline: 16298775]
41. Horne J, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4(2):97-110. [Medline: 1027738]
42. Randler C. Morningness-eveningness comparison in adolescents from different countries around the world. Chronobiol Int 2008 Nov;25(6):1017-1028. [doi: 10.1080/07420520802551519] [Medline: 19005902]
43. Randler C. Gender differences in morningness–eveningness assessed by self-report questionnaires: A meta-analysis. Personality and Individual Differences 2007 Nov;43(7):1667-1675 [FREE Full text] [doi: 10.1016/j.paid.2007.05.004]
44. Tonetti L, Fabbri M, Natale V. Sex difference in sleep-time preference and sleep need: a cross-sectional survey among Italian pre-adolescents, adolescents, and adults. Chronobiol Int 2008 Sep;25(5):745-759. [doi: 10.1080/07420520802394191] [Medline: 18780201]
45. Rahafar A, Randler C, Díaz-Morales JF, Kasaeian A, Heidari Z. Cross-cultural validity of Morningness-Eveningness Stability Scale improved (MESSi) in Iran, Spain and Germany. Chronobiol Int 2017;34(2):273-279. [doi: 10.1080/07420528.2016.1267187] [Medline: 28055262]
46. Merikanto I, Kronholm E, Peltonen M, Laatikainen T, Lahti T, Partonen T. Relation of chronotype to sleep complaints in the general Finnish population. Chronobiol Int 2012 Apr;29(3):311-317. [doi: 10.3109/07420528.2012.655870] [Medline: 22390244]
47. Duarte L, Menna-Barreto L, Miguel M, Louzada F, Araújo J, Alam M, et al. Chronotype ontogeny related to gender. Braz J Med Biol Res 2014 Apr;47(4):316-320 [FREE Full text] [Medline: 24714814]
48. Kohyama J. Self-Reported Academic Performance and Lifestyle Habits of School Children in Japan. Int J Child Health Nutr 2017 Sep 11;6(3):90-97. [doi: 10.6000/1929-4247.2017.06.03.1]
49. Polugrudov AS, Panev AS, Smirnov VV, Paderin NM, Borisenkov MF, Popov SV. Wrist temperature and cortisol awakening response in humans with social jetlag in the North. Chronobiol Int 2016;33(7):802-809. [doi: 10.3109/07420528.2016.1168829] [Medline: 27101215]
50. Johnsen MT, Wynn R, Bratlid T. Optimal sleep duration in the subarctic with respect to obesity risk is 8-9 hours. PLoS One 2013;8(2):e56756 [FREE Full text] [doi: 10.1371/journal.pone.0056756] [Medline: 23457611]
51. Farkova E, Mankova D, Koprivova J. Is increased body mass index associated with larger social jet-lag? Sleep Medicine 2017 Dec;40:e93. [doi: 10.1016/j.sleep.2017.11.270]
52. Yamada T, Shojima N, Yamauchi T, Kadowaki T. J-curve relation between daytime nap duration and type 2 diabetes or metabolic syndrome: A dose-response meta-analysis. Sci Rep 2016 Dec 02;6:38075 [FREE Full text] [doi: 10.1038/srep38075] [Medline: 27909305]
53. Yamada T, Hara K, Shojima N, Yamauchi T, Kadowaki T. Daytime Napping and the Risk of Cardiovascular Disease and All-Cause Mortality: A Prospective Study and Dose-Response Meta-Analysis. Sleep 2015 Dec 01;38(12):1945-1953 [FREE Full text] [doi: 10.5665/sleep.5246] [Medline: 26158892]
54. Guo VY, Cao B, Wong CKH, Yu EYT. The association between daytime napping and risk of diabetes: a systematic review and meta-analysis of observational studies. Sleep Med 2017 Sep;37:105-112. [doi: 10.1016/j.sleep.2017.01.018] [Medline: 28899519]
55. Malone SK, Patterson F, Lu Y, Lozano A, Hanlon A. Ethnic differences in sleep duration and morning-evening type in a population sample. Chronobiol Int 2016;33(1):10-21 [FREE Full text] [doi: 10.3109/07420528.2015.1107729] [Medline: 26654569]
56. Randler C, Prokop P, Sahu S, Haldar P. Cross-cultural comparison of seven morningness and sleep-wake measures from Germany, India and Slovakia. Int J Psychol 2015 Aug;50(4):279-287. [doi: 10.1002/ijop.12098] [Medline: 25175755]

Abbreviations

BMI: body mass index
CPC: cardiopulmonary coupling
MCTQ: Munich ChronoType Questionnaire

XSL•FO
RenderX

MEQ: Morningness-Eveningness Questionnaire
MSFsc: adjusted mid-point of sleep on free days
SJL: social jetlag

Edited by G Eysenbach; submitted 24.01.19; peer-reviewed by P Innominato, J Di, C Fernández; comments to author 27.02.19; revised version received 27.03.19; accepted 29.03.19; published 11.05.19.

Please cite as:
Zhang Z, Cajochen C, Khatami R
Social Jetlag and Chronotypes in the Chinese Population: Analysis of Data Recorded by Wearable Devices
J Med Internet Res 2019;21(6):e13482
URL: https://www.jmir.org/2019/6/e13482/
doi:10.2196/13482
PMID:31199292

©Zhongxing Zhang, Christian Cajochen, Ramin Khatami. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 11.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Machine Learning Approach for the Detection and Characterization of Illicit Drug Dealers on Instagram: Model Evaluation Study

Jiawei Li[1,2], MS; Qing Xu[1,2], MAS; Neal Shah[2], BS; Tim K Mackey[1,2,3,4], MAS, PhD

[1]Department of Healthcare Research and Policy, University of California - San Diego, Extension, La Jolla, CA, United States

[2]Global Health Policy Institute, La Jolla, CA, United States

[3]Department of Anesthesiology, University of California - San Diego, School of Medicine, La Jolla, CA, United States

[4]Division of Infectious Disease and Global Public Health, University of California - San Diego, School of Medicine, La Jolla, CA, United States

Corresponding Author:
Tim K Mackey, MAS, PhD
Department of Anesthesiology
University of California - San Diego, School of Medicine
8950 Villa La Jolla Drive
A124
La Jolla, CA, 92037
United States
Phone: 1 9514914161
Email: tmackey@ucsd.edu

Abstract

Background: Social media use is now ubiquitous, but the growth in social media communications has also made it a convenient digital platform for drug dealers selling controlled substances, opioids, and other illicit drugs. Previous studies and news investigations have reported the use of popular social media platforms as conduits for opioid sales. This study uses deep learning to detect illicit drug dealing on the image and video sharing platform Instagram.

Objective: The aim of this study was to develop and evaluate a machine learning approach to detect Instagram posts related to illegal internet drug dealing.

Methods: In this paper, we describe an approach to detect drug dealers by using a deep learning model on Instagram. We collected Instagram posts using a Web scraper between July 2018 and October 2018 and then compared our deep learning model against 3 different machine learning models (eg, random forest, decision tree, and support vector machine) to assess the performance and accuracy of the model. For our deep learning model, we used the long short-term memory unit in the recurrent neural network to learn the pattern of the text of drug dealing posts. We also manually annotated all posts collected to evaluate our model performance and to characterize drug selling conversations.

Results: From the 12,857 posts we collected, we detected 1228 drug dealer posts comprising 267 unique users. We used cross-validation to evaluate the 4 models, with our deep learning model reaching 95% on F1 score and performing better than the other 3 models. We also found that by removing the hashtags in the text, the model had better performance. Detected posts contained hashtags related to several drugs, including the controlled substance Xanax (1078/1228, 87.78%), oxycodone/OxyContin (321/1228, 26.14%), and illicit drugs lysergic acid diethylamide (213/1228, 17.34%) and 3,4-methylenedioxy-methamphetamine (94/1228, 7.65%). We also observed the use of communication applications for suspected drug trading through user comments.

Conclusions: Our approach using a combination of Web scraping and deep learning was able to detect illegal online drug sellers on Instagram, with high accuracy. Despite increased scrutiny by regulators and policymakers, the Instagram platform continues to host posts from drug dealers, in violation of federal law. Further action needs to be taken to ensure the safety of social media communities and help put an end to this illicit digital channel of sourcing.

(J Med Internet Res 2019;21(6):e13803) doi:10.2196/13803

KEYWORDS

opioids; social media; narcotics; substance abuse; machine learning; internet; prescription drug abuse; artificial intelligence

XSL•FO
RenderX

Introduction

Background

In June 2018, the US Food and Drug Administration (FDA) held the *Online Opioid Summit*, a 1-day meeting seeking to generate momentum around the need to combat illicit internet sales of opioids [1]. In addition to federal agencies, several internet and social media companies were in attendance, including Google (which operates YouTube and Google+), Twitter, Facebook (which operates Instagram and WhatsApp), Pinterest, and other e-commerce, technology, and patient safety organizations [2]. The meeting was organized to facilitate cooperation among these stakeholders to address illicit internet opioid sourcing and diversion, a challenge that adds fuel to a national public health emergency that claims the lives of an average of 130 people daily from addiction-related overdose [3].

Importantly, federal law explicitly prohibits the internet sale of controlled substances as enforced by the 2008 Ryan Haight Online Pharmacy Consumer Protection Act (RHA) [4,5]. Named after a Californian adolescent who died in 2001 after overdosing on Vicodin purchased from an online drug seller without a prescription, the RHA was meant to curb the use of the internet as an alternative and convenient channel of sourcing [6]. However, since Mr Haight's death, the internet ecosystem has rapidly proliferated and diversified, now populated by illegal internet pharmacies, social media posts from illegal sellers, and dark Web vendors, all who have been implicated in illegal online opioid sales [6-12].

Though illegal prescription drug sales are often found by users through search engine results and internet pharmacy advertisements (including spam email), popular social media platforms have emerged as a direct-to-consumer marketing tool for illegal sellers [2,12-15]. Previous research and investigative reporting have detected illegal opioid sales and drug dealing on several social media platforms, including Twitter, Facebook, and Instagram [7,12,16-21]. As Twitter provides a convenient way of accessing data through its application programming interface (API), many substance abuse infoveillance studies have focused on this microblogging platform [22-25]. Our own prior studies used both supervised machine learning classifiers and an unsupervised topic model to detect internet pharmacies selling opioids (including fentanyl) [6,7,21,26]. Others have primarily focused on analyzing Twitter messages for opioid and substance abuse behavior with manually annotated data, examining social circles of users, measuring user sentiment, using natural language processing, and using deep learning [22,23,27,28]. Other studies have used deep learning models to detect and describe adverse drug reactions via Twitter [29,30].

However, there are far fewer studies that have conducted infoveillance research on the Instagram platform, likely because of the difficulty of collecting Instagram data and the different data features that require additional data cleaning and processing. Instagram is an image and video sharing social media platform (reaching 1 billion monthly users in 2018) and is particularly popular among young adults (ie, 71% of those aged between 18 and 24 years), a critical demographic for substance use initiation [17]. Prior studies by Zhou et al have analyzed Instagram data primarily for substance abuse behavior and did not use deep learning models but instead used other machine learning approaches [28]. One relevant study by Yang and Luo used a model based on multitask learning that analyzed both images and text to track and classify drug abuse–related posts, including differentiating for drug dealers [17]. The study analyzed the user timelines of identified posts to differentiate drug dealers from users who exhibited drug use behavior and achieved a high classification accuracy of 88% [17].

Objective

Building on these prior studies that have used different big data and machine learning approaches to detect substance abuse behavior and illegal drug selling on social media, this study describes the use and evaluation of a deep learning model to better automate the detection of illegal opioid and other illicit drug sales on Instagram. Our study focuses on detecting illegal drug selling posts (not accounts) from hashtag searches and using deep learning to analyze text from posts.

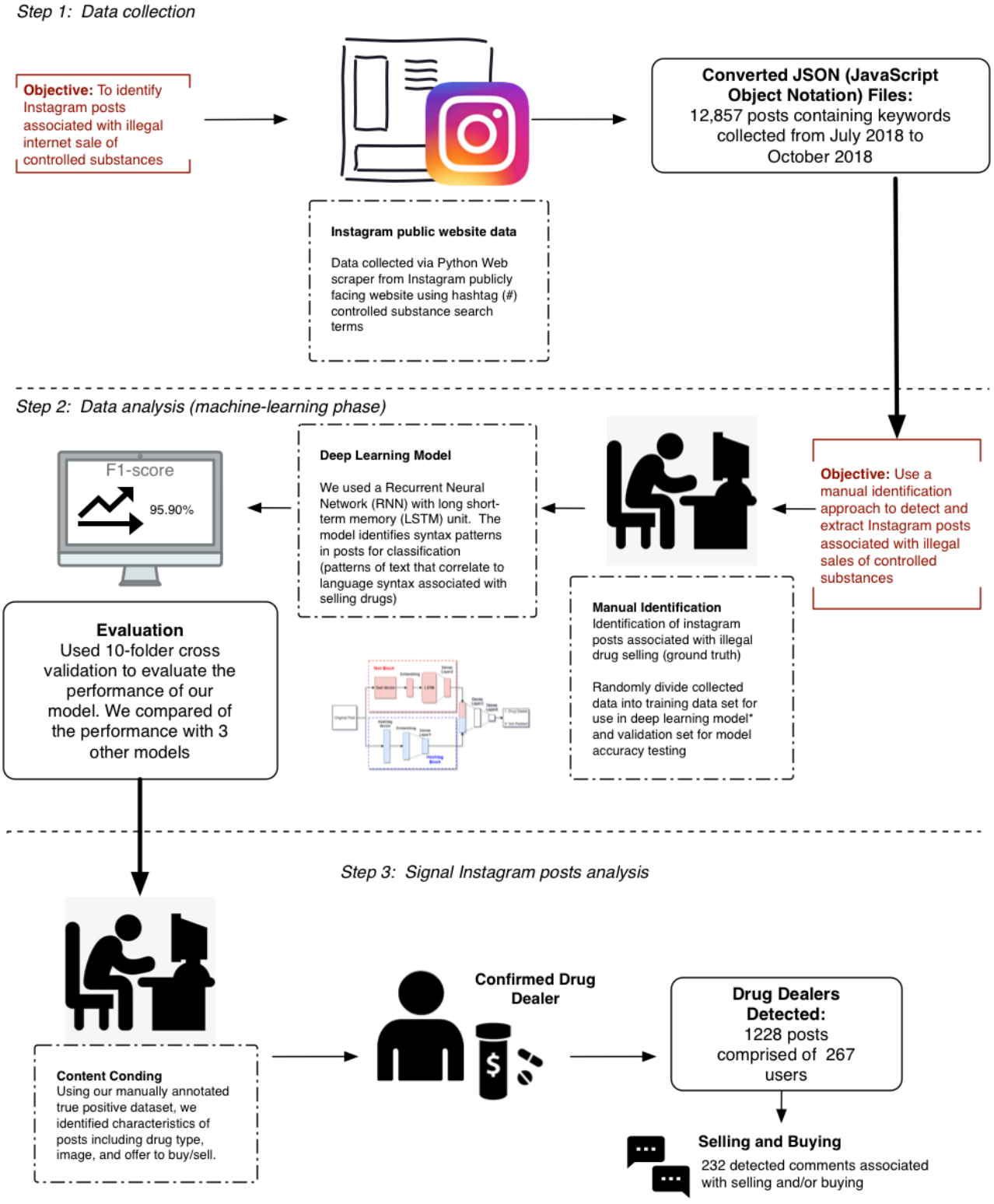
Methods

Overview

Our study comprised 3 phases: data collection, data processing, and model training. The goal of our study was to develop and evaluate a machine learning approach that has the best performance for identifying illicit opioid drug selling via Instagram (Facebook, Inc.). To accomplish this, we first collected a set of posts that contained suspicious drug selling behavior by conducting automated searches for opioid-related hashtags and posts. We then used posts detected in these hashtag searches as a training set for our deep learning model, so we could better enable detection of posts in the entire corpus of all messages collected (see Figure 1 for summary of methods). Importantly, Instagram is a platform that has different ways of presenting messages from its users. For other social media platforms such as Twitter and Facebook, the main content in the post is often the text accompanied by hashtags. However, on the basis of the method of searching for messages on Instagram (ie, users search for posts and topics by hashtags), the more hashtags a post contains, the easier it will be found. Therefore, a post for Instagram usually comprises 3 main pieces of content: text, hashtags, and an image. In this study, we will analyze the performance of models examining text and hashtags to determine what is the best approach to detecting illegal drug dealers.

XSL•FO
RenderX

Figure 1. Summary of study methodology.



Data Collection Phase

Since July 2018, Instagram has disabled many functions of its public API and has set limitations on data collection. To collect a more representative dataset than what is available from its limited public API, we developed a method to Web scrape results from the Instagram platform website based on opioid keyword hashtags (a word or phrase preceded by a hash character, #) search results. Our Web scraper was built in the computer programming language Python and converted source code from the Instagram website into JavaScript Object Notation (JSON) data files. Data scraped from the website included the text of the Instagram post, comments to the post by Instagram users, and associated metadata (eg, date, time, and user profile information).

As search on Instagram is conducted using hashtags (eg, #subject), we chose an initial list of hashtags related to controlled substances identified via manual searches and then

XSL•FO
RenderX

populated these opioid-related hashtags into our automated search and Web scraping data collection process. This allowed us to discover a greater number of opioid and drug-related hashtags that are present in the Instagram community. Our initial set of hashtag keywords included *xanaxangel*, *percocert* (Percocet misspelled), *adderrall* (Adderall misspelled), and *hydrocodoneacetaminophen* (misspelled.) These keywords relate to hashtags associated with common controlled substance drug names that were used on Instagram at the time of the study. Many of these keywords are misspelled as Facebook and Instagram currently disables search results for certain explicit opioid keywords in search queries [2]. However, alternative opioid hashtag keywords are relatively easy to find, including some that are derived from the platform's own suggested alternative hashtags when conducting searches [31].

Importantly, hashtags used by drug dealers are different depending on what type of drug(s) they are selling. Hence, it is important to expand the number and diversity of possible drug-related hashtags to collect a better sample of data to analyze. To increase the number of hashtags likely related to drug dealing, we examined search results for our initial set of hashtags using a 2-loop process during our automated search and Web scraping. In the first loop, we captured data from an individual post under a certain hashtag in the initial set of hashtags used. Our Web scraper continued to collect the JSON data from the source code for each of these posts until it reached a set limitation (in this study we ended our search loop when the post was older than 3 months). We then filtered out the hashtags from each post and chose other hashtags that contained keywords associated with controlled substances and illicit drugs by manually inspecting all hashtags collected in the loop. Hashtags that contained opioid and drug-related keywords or combinations thereof were then added to the hashtag list so they could be searched in a second loop that would go through all the hashtags identified. A full list of hashtags identified in this study is available in Multimedia Appendix 1.

Data collection occurred from July to October 2018, with hashtags limited to English language and with no set geographic or other filters for posts. In our data cleaning process, we also discarded duplicate posts that were replicated when collecting posts under different hashtags search loops, and then removed hashtags and stop words before textual analysis in the data processing phase described below.

Data Processing Phase

An Instagram post generally contains content with the user's text (ie, message) and also hashtags to self-curate the content and associate it with other user posts and comments. Unlike normal text, hashtags can be placed randomly in the post, which could affect the accuracy of the machine learning model used in this study, as our deep learning model is based on learning the pattern of the text. To address this, we took the following additional steps to process our data:

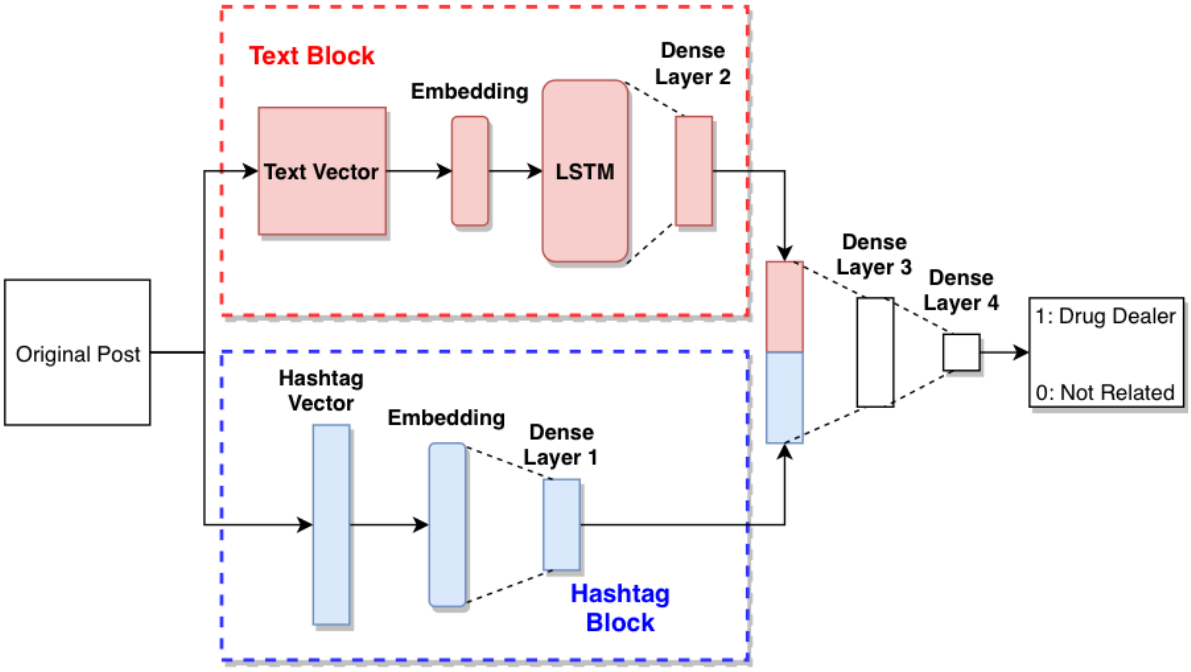
1. We eliminated duplicate results of posts that appeared in multiple hashtag searches.
2. From each post, we extracted the text and removed any hyperlinks, special characters (eg, emoji, !, @), and stop words (eg, hers, between, yourself). We used the Natural Language Toolkit package in Python to remove stop words. We did not exclude # in this study, as # represents the hashtags of keywords that we specifically wanted to analyze.
3. We then built a dictionary based on the words in all texts; each word had a corresponding index. Then we transformed the word into an index in each text.
4. We kept text that had more than 1 index after step 2 as original data; then we removed all the hashtags from the original data and removed the texts that had no remaining words left. These data are referred to as no hashtag data.
5. We eliminated the duplicate texts in each dataset after step 3. Hence, all the texts remaining were unique with a different pattern.
6. We then manually annotated all the posts detected (including text and images) to identify and classify posts that involved illegal drug dealing or selling to establish ground truth for model evaluation. We accomplished this by using a binary coding scheme of yes or no based on assessing whether a post contained text or image information about a prescription opioid, controlled substance, or other suspected illicit drug product and that the post also included contact information on how to trade or purchase the drug from the dealer. The second and the third author coded posts independently and achieved a high intercoder reliability for results (kappa=0.98). For inconsistent results, both authors met and reviewed the posts together with the last author (a subject matter expert in internet substance abuse behavior) and all authors conferred on the correct classification of the post.
7. Using the manually annotated ground truth data in step 6, we then evaluated the performance of a deep learning model to identify illegal drug sellers by comparing it with other machine learning models.

Model Training

In this study, we used 4 supervised models to analyze our Instagram data: decision tree, random forest (RF), support vector machine (SVM), and a deep learning model we developed [32-34]. The first 3 models are traditional machine learning models, which perform well for classification tasks [35-37]. For this study's deep learning model, we used a recurrent neural network with long short-term memory (LSTM) unit to study the pattern of text in a post. This model is well-suited for classifying, processing, and making predictions based on time series data and also exhibits high performance on speech text analysis [38-41].

XSL•FO
RenderX

Figure 2. Structure of deep learning model. Embedding layer: input_dim is 29,832, which is the size of the dictionary; the input_length for text is 50, input_length for hashtag is 15; the output_dim is 400. Long short-term memory layer: contains 800 units, with dropout=0.2, recurrent_dropout=0.2. Dense layer 1: unit=200, activation=sigmoid. Dense layer 2: unit=200, activation=sigmoid. Dense layer 3: unit=200, activation=sigmoid. Dense layer 4: unit=1, activation=sigmoid. Optimizer: Adam (learning rate set at 0.0001). Loss: Binary_crossentropy. LSTM: long short-term memory.



The structure of the model is shown in Figure 2. It contains 2 major parts, text block and hashtag block. The text block was used to store the information in the text only, the hashtag block was used to store the information from the hashtags only. When we input the data, we separated the original text into the text part and hashtags part, and for each part we generated a vector based on the dictionary we built in the data processing stage. The length of the text vector was 50 words, which is the average length of texts in our dataset (if the size of the vector is too large, it will increase the workload for the model; if it is too short, we may lose signal data). If the text was longer than 50 words, we only chose the first 50 words to maintain the efficiency of the model; if it was shorter, we filled the rest of the vector with 0. Then we used word embedding to generate the text into a word vector with dimensions (50, 400). The strategy for hashtags was the same, but the average length of hashtags was set to 15 characters.

The text block contains 1 layer of LSTM combined with 1 fully connected network (FCN) [42]. LSTM is a recurrent neural network that can be used to predict the next word based on the current word. Hence, it will learn the correlation between different words, which is the pattern of the text, during the training process. The FCN (dense layer) is used to learn features from the original data and to keep as much information as possible in another vector (in our study we kept the output from the former layer into a smaller dimension vector). The dense layers 1 and 2 in these 2 blocks will keep the signal from text to a much smaller vector. We then merged these 2 vectors together and used another 2 FCNs (dense layers 3 and 4) to generate the prediction.

Every time the model gives a prediction (possibility for 1=drug dealer and 0=not related; if the possibility is larger than 0.5, we consider it as drug dealer), the model will calculate the difference between the prediction and the label; the difference is referred to as the loss. If the loss is low, that means the model is performing well on prediction. There are many ways to calculate the loss; in our model, we used binary cross-entropy loss [43]:

$$\text{Loss}=-(y\times\log(p)+(1-y)\times\log(1-p))$$

This is the format for a binary class only; y represents the ground truth of each text (either 0=not drug dealer or 1=drug dealer), p represents the predicted probability that the text is of the class. The loss shows the difference between our prediction and ground truth.

After we get the loss, the model will run back propagation to update the parameters in LSTM and FCN to decrease the loss value [44]. The updating is based on the learning rate and optimizer we choose. The learning rate is used to keep the learning process in an acceptable range. If the learning rate is too small, it will take a longer time for the model to achieve the best parameters; if it is too large, it may never reach the local minimum for the loss. Therefore, an Adam optimizer algorithm was used for back propagation, which allows the learning rate to adapt based on the parameters which will make large updates for infrequent parameters and small updates for frequent parameters [45].

During the updating process, the model will learn how to distinguish what a drug dealer's post looks like. On the basis of this pattern, the model can identify syntax patterns in posts to classify them as signal posts (patterns of text that correlate to language syntax associated with selling drugs) and can then separate posts that contain relevant hashtag keywords, whose syntax is not related to drug selling.

XSL•FO
RenderX

The data were separated into a training set and validation set; the training set was used to update the model and the validation set was used to evaluate the performance of our model, which we discuss further in the Model Evaluation section. In the training process, the deep learning model will keep looping on learning the training set to reach the minimum of the loss. However, this process could cause overfit, which will make the model too sensitive to the pattern of the texts in the training set. To prevent overfit, we need another dataset that is different from the training set (ie, the validation set). If the loss of the validation set is increasing, it means that the model learned too much from the training set and the training needs to be stopped. In our model, we compare the most recent validation loss with all 4 previous validation losses; if it is larger (the loss is continually increasing), the model will stop training.

Ethics Approval and Consent to Participate

Ethics approval and consent to participate were not required for this study as data derived for this study were available from public sources and there were no interactions with social media users.

Data Availability

The pretrained model for this study is available on GitHub under the user account name Mathison under the file Instagram_drugdealer_detection.

Results

Model Evaluation

We collected a total of 12,857 Instagram posts over a 3-month period (July 19, 2018 to October 18, 2018) that we used for analysis. There were a total of 1228 drug dealer posts based on the manual annotation we conducted, which comprised 267 unique users. As of October 18, 2018, 206 of these posts were still active and viewable on the Instagram platform. As the volume of the target posts is low compared with the total dataset, we used 10-folder cross-validation to evaluate the performance of our model. This method is used to estimate how the model is expected to perform when facing limited samples. We shuffled each dataset (original data and no hashtag data) randomly and separated them into 10 folders—9 of them are the training set (in the deep learning model we used 70% as the training set and 30% as the validation set) and the rest are the test set used to assess the area under the curve (AUC), precision, recall, and the F1 score of the model. When calculating the AUC, we used the whole dataset rather than the cross-validation. There were a total of 10 iterations; for each iteration, we chose a different folder as the test set. Hence, this allowed us to ensure that each text could be used to make prediction. For each iteration, the model was reset. After we finished all iterations, we could calculate the average score for each model.

For the deep learning model, to prevent the problem of overfit, we separated the original training set into 2 parts, 70% of the set as the training set and 30% as the validation set. This allowed us to ensure that the model we generate from each iteration can have the best performance on the test set.

The results from our model evaluation are separated into 3 parts according to the data we preprocessed (Table 1): text with hashtags, hashtags only, and text without hashtags.

On the basis of this evaluation, the deep learning model has the best performance compared with the other 3 models based on the F1 score. However, the precision of the deep learning model does not show a better result than RF or SVM, suggesting that deep learning is not more effective at filtering for false positives.

Table 1. Performance for each model based on variations of text and hashtag use.

Performance measure	Decision tree	Random forest	Support vector machine	Study model
Text with hashtags, %				
Precision	95.05	96.00	96.86	94.81
Recall	82.15	86.08	81.21	91.42
F1 score	88.13	90.77	88.35	93.09
Area under the curve	96.67	96.85	97.18	98.12
Hashtags only, %				
Precision	86.22	94.14	95.39	89.60
Recall	86.50	87.13	84.24	88.89
F1 score	86.36	90.50	89.47	89.24
Area under the curve	95.95	95.23	95.43	94.32
Text without hashtags, %				
Precision	88.49	97.07	97.80	93.60
Recall	93.08	91.31	89.32	98.31
F1 score	90.73	94.11	93.37	95.90
Area under the curve	95.56	94.85	93.49	99.12

XSL•FO
RenderX

The possible reason is that drug dealers on Instagram usually have similar distinct text formats of selling drugs, which includes providing contact information with a phone number or email address. Therefore, once the pattern of the text is established, the deep learning model can identify these posts with greater ease. This is also shown in the results; the recall of the deep learning model is much higher than in other models. However, the text patterns of other nonrelated posts are not similar, so the deep learning model may be confused by random patterns which makes its precision, compared with other models, lower.

Table 1 shows that the model performance on hashtags only does not improve when compared with performance on texts with hashtags, especially for the deep learning model. However, when we use the text without the hashtags, all the scores increased except for the recall of the deep learning model compared with performance on texts with hashtags (Table 1). These performance results suggest that removing hashtags can increase the accuracy of the model, but may increase false positives for the deep learning model. This may occur when the user is engaging in nondrug related sales, but uses the same text pattern with different hashtags than drug dealers, as removal of hashtags makes these texts patterns appear more similar to posts from actual drug dealers.

Text Analysis of Drug Dealer Posts

When manually annotating the text contained in the 1228 drug selling posts, we identified 2 forms of communication by users: (1) the use of hashtags (#) in front of illicit drug-related keywords to self-curate content and (2) descriptive language of drug selling–related activity in combination with related images posted by users. Each post can contain a combination of types of text (eg, hashtags and/or descriptive text), images or videos, and other metadata associated with the Instagram user. The vast majority of posts (1196/1228, 97.39%) had hashtags, whereas 32 out of 1228 posts (2.61%) only had descriptive language with no hashtags, which we suspect were posts modified after data collection occurred and before manual annotation.

The majority of hashtags detected in signal posts were related to the controlled substance Xanax (Alprazolam, a nonopioid controlled substance; 1078/1228 posts, 87.78%), followed by oxycodone/OxyContin (321/1228, 26.14%) and illicit drugs lysergic acid diethylamide (LSD; 213/1228, 17.34%) and 3,4-methylenedioxymethamphetamine (MDMA; 94/1228, 7.65%; see Table 2 for summary).

Table 2. Number of posts related to controlled substance hashtags (N=1228).

Drug name[a]	Hashtag[b]	Posts, n (%)[c]
Xanax	#xanax	802 (65.3)
	#xanaxfamily	530 (43.1)
	#2mgxanax	321 (26.1)
	#zanax	112 (9.1)
	#greenxanax	84 (6.8)
	Total	1078 (87.8)
Oxycodone/OxyContin	#oxycodone	30 (2.4)
	#oxycodine	261 (21.3)
	#oxy80s	213 (17.3)
	#oxycontin	215 (17.5)
	#oxicotin	233 (18.9)
	#oxicodone	212 (17.2)
	Total	321 (26.1)
Lysergic acid diethylamide	#LSD25	138 (11.2)
	#LSDtabs	130 (10.5)
	Total	213 (17.3)
3,4-Methylenedioxy-methamphetamine	#mdmapills	50 (4.1)
	#mdmaforsale	21 (1.7)
	#mdmazing	40 (3.3)
	#mdmaonline	21 (1.7)
	Total	94 (7.6)

[a]Drug name column relates to drug detected in the image and text of the post.

[b]Hashtag refers to the presence of a hashtag in a post detected.

[c]Posts is the number of posts with the hashtag and the percentage of total posts that contained the hashtag.

XSL•FO
RenderX

When manually annotating posts that contained hashtags, we found that the use of hashtags can be categorized into 3 major groups: (1) hashtags with controlled substance drug names and other illicit drug names, (2) slang or other codewords used in the Instagram user community for specific controlled substances and substance use behavior, and (3) other keywords and codewords describing selling and other promotional behavior (eg, shipping and selling). Examples for each of these categories are provided below:

- Drug name: Generic or brand name of a drug in a hashtag (eg, #xanax, #oxycodone, #oxycontin, #LSD, #MDMA)
- Codewords: Codewords for controlled substances, including (1) misspelled drug name, for example, "#zanax", (2) extended drug name, for example, "#mdmaforsale", "#2mgxanax", and (3) street name of drug, for example, "#whitebar"
- Keywords related to sale or shipping, for example, "#forsale", "#shipping", etc

When assessing posts with descriptive text describing actions or behaviors of drug dealers, we were able to classify posts into 2 additional categories: (1) drug sale promoting language, for example, "interested in placing order without prescription", "order now for quick delivery" and (2) contact information of purported drug seller. On the basis of these 2 categories, Instagram drug dealers appear to clearly indicate in their descriptive text of their posts an offer for sale and also provide other users with information on how to contact the seller. Contact information generally included an email address, phone number, or user account information about a communication-based application or mobile app. Among posts containing communication applications or apps, Wickr (445/1228, 36.23%), Telegram (245/1228, 19.95%), Kik app (225/1228, 18.32%), and WhatsApp (188/1228, 15.30%) were utilized. In some cases, a drug seller might also include descriptive text in the post referring to their account profile to reference contact information instead of providing it in the text of the post.

From the 1228 detected posts we analyzed, 232 of these posts explicitly included an offer for sale and offer to buy by the users. In this case, there was an Instagram post or comments within a post from a user offering to sell drug(s) (with contact information) and a comment from another user that asked for more information or offered to buy the drug.

Manual Image Annotation of Drug Dealer Posts

In addition to text analysis, we also collected a total of 260 pictures from the 1228 posts analyzed. These images were manually annotated to determine if they were related to controlled substances or illicit drugs (primarily coded for whether they included a picture of a controlled substance or suspected illicit drug product) or if they were unrelated images. A total of 252 images (252/260, 96.9%) included pictures or images of different types of drugs that can be categorized into 5 different categories.

In the first category which included prescription-controlled substances only, there were 175 posts (175/260, 67.3%) comprised of the following controlled substances: Xanax (41 posts), Alprazolam (34 posts), oxycodone/OxyContin (25 posts), Adderall (17 posts), and amphetamine and dextroamphetamine (14 posts). In the second category that contained images of illicit drugs only, there were 60 images comprising LSD (19 posts, including blotter paper soaked in LSD), ecstasy/MDMA (14 posts), cannabis (13 posts), and magic mushrooms (7 posts; see Figure 3).

In the third category, there were 15 images that contained both a prescription drug and illicit drug. In the fourth category, there were 5 images we could not classify, but which we suspected as either illicit drugs or other drug-related manufacturing materials (see Figure 4.)

Finally, in the fifth category, separate from images of drug products, we also detected images that included typed or written information on a physical medium communicating the drug dealers' contact information (see Figure 5).

XSL•FO
RenderX

Figure 3. Examples of Instagram posts of illegal drug sale categories (user information and text from post have removed). (1) A post of prescription drugs; (2) a post of lysergic acid diethylamide (LSD); (3) a post with written contact information imbedded in the image; and (4) post with multiple drug types.

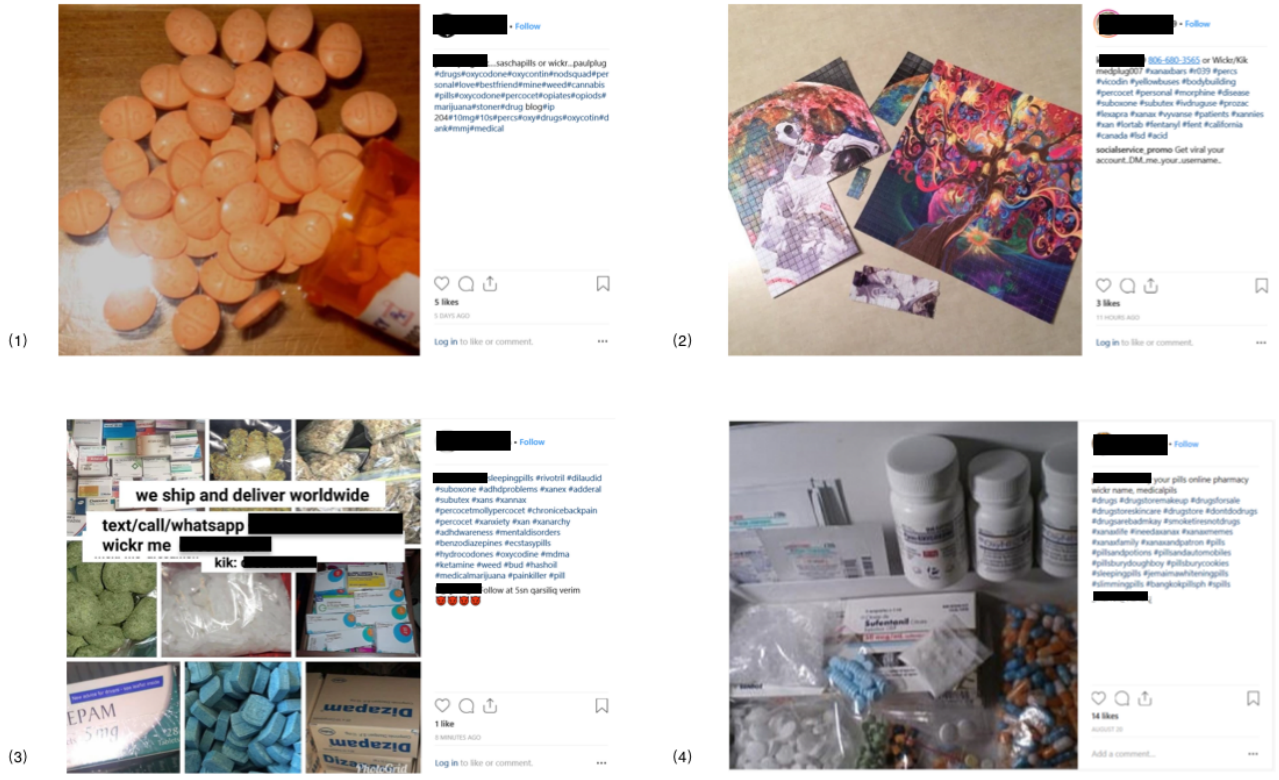
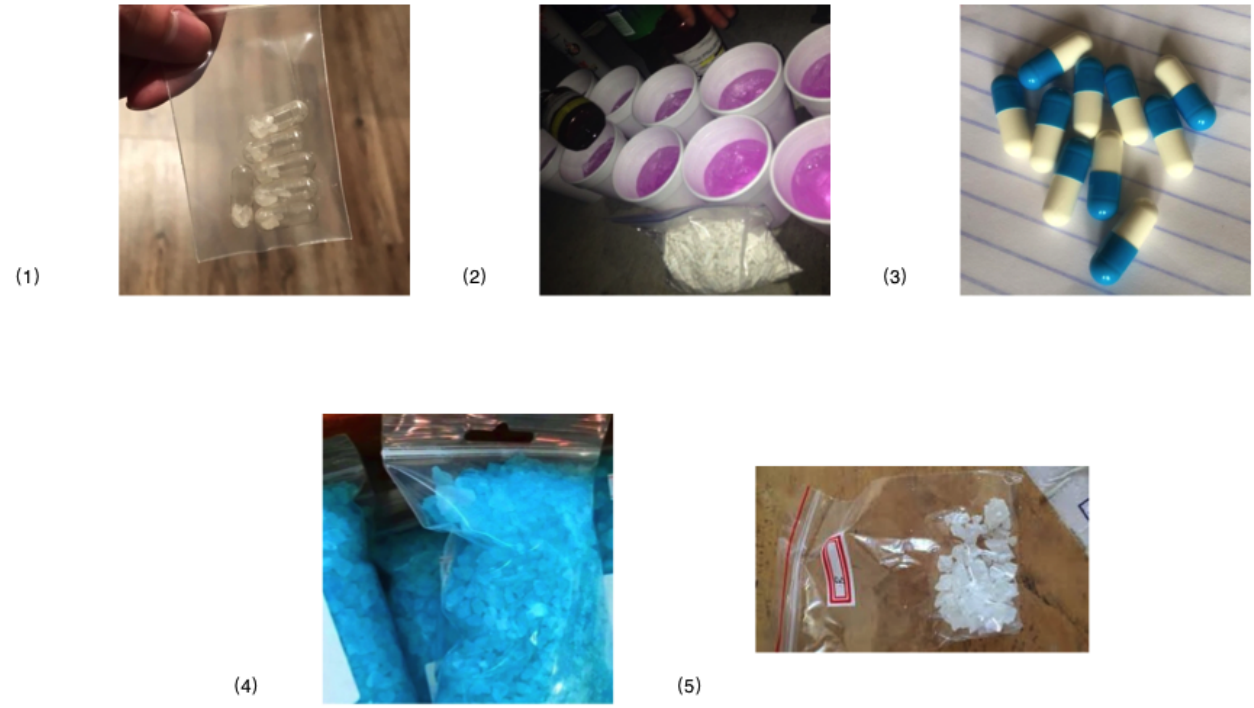
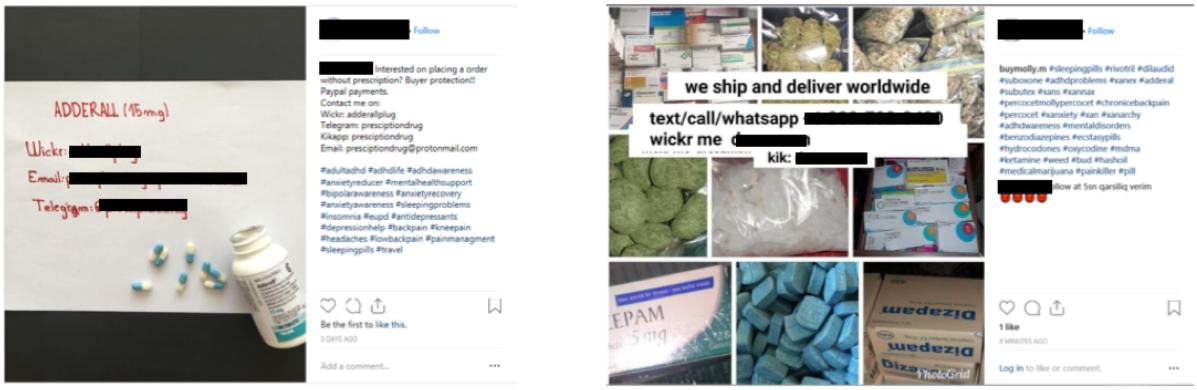


Figure 4. Examples of Instagram posts with suspected drug products (user information removed). The 5 unclassified pictures include (1) clear capsules with white crystalline granules, (2) cups with pink liquid, (3) blue and white capsules with no drug identification, (4) plastic bags of blue crystals, and (5) a bag of white crystals with a label “B”.



XSL•FO
RenderX

Figure 5. Examples of Instagram posts with written contact information. There were 60 images that included either typed or written contact information.



Discussion

Principal Findings

Over a 3-month period, we used a combination of hashtag keyword searches, Web scraping, deep learning, and manual annotation to detect and characterize 1228 posts from 267 unique users that we suspected were associated with illegal drug dealing on Instagram. Several prescription-controlled substances, illicit drugs, and other suspected drug products were detected as being offered for sale purportedly by drug sellers/dealers. In addition, we observed Instagram users having conversations via comments purporting to both offer a sale of a drug product and receiving an offer to buy from another user (ie, generally a comment from a seller to another user to contact them directly via a third-party communication application to enter into a drug-related transaction.)

These initial results are alarming and generally conform to existing nonscientific investigational news reports on the subject, in addition to published research on illegal drug sales on this and other social media platforms such as Twitter [6,7,17]. As a result of emerging evidence linking the risks of illicit drug access and diversion via social media technology there has been increased public attention and scrutiny. Congressional bodies, including the House Energy and Commerce Committee, have called on social media company executives to explain why their platforms facilitate this illegal activity [46,47]. This includes sharp inquiry from Congressman David McKinley (West Virginia), whose state has been heavily impacted by the opioid epidemic. Mr McKinley questioned both Twitter co-founder and Chief Executive Officer Jack Dorsey and Facebook Chairman and Chief Executive Officer Mark Zuckerberg in 2018 congressional testimony on what steps their platforms were taking to remove illegal opioid sellers and how they will protect the public [47,48]. Mr Zuckerberg responded that the enormity of data makes it hard to monitor content and that *artificial intelligence* approaches to *proactively* find content were needed [47].

In April 2018, Facebook and Instagram took action by blocking opioid-related hashtag searches on their platform and reportedly suspending accounts [16,20]. However, our study indicates that drug sellers continue to populate Instagram despite these actions and that these communities have changed their use of hashtags possibly to avoid detection. Hence, to carry out the legislative intent of the RHA to promote patient safety and prevent substance abuse behavior, there is clear need for innovative technology solutions that have high accuracy and are scalable and can help all parties (including technology companies, regulators, and law enforcement) detect, classify, and take action against digital drug dealers.

Study Limitations

Our study has certain limitations. First, this study was limited to a short period of data collection and we did not purposely sample Instagram accounts or consider user characteristics such as age, gender, or other demographics. Hence, study results are not generalizable and are not necessarily representative of the Instagram user community. In addition, demographic data are not always readily available in the metadata or the user account information or the post, and if available, may not be accurate. Future studies should assess which specific user communities may be at higher risk for illegal drug sourcing online. We also did not engage with users or verify if drugs purportedly being sold were actually available or sold to other users. Given this limitation, we cannot say with certainty that these drugs were actually being sold. Conducting test purchases of controlled substances and other illicit drugs is prohibited by federal law. However, drug dealers often post pictures of drug products to demonstrate to users that they have availability and we did not observe comments reporting scams or failed drug buys. We also did not use multimodal or synchronous approaches to develop a classification model based on both text and images as used by Yang and Luo, an approach that could improve performance of the model and should be explored in future studies. In addition, though data were collected and analyzed within 3 months of collection, we relied on manual annotation to establish validity of results and evaluate the performance of our model. This time lag because of manual annotation may have resulted in some posts being removed or modified before analysis or manual annotation. Specifically, drug dealers may self-delete posts after they have completed a transaction. Future studies should continue the iterative process of establishing training datasets to inform machine learning models that can more quickly and accurately detect illicit drug dealing.

XSL•FO
RenderX

Conclusions

In this study, we evaluated a deep learning model to detect drug dealers on Instgram. The deep learning model based on LSTM performed the best compared with the other 3 models evaluated. Furthermore, we compared the deep learning model's performance with hashtags against messages with only text in Instagram posts and demonstrated that the model yields better results from text without the hashtags, despite the risk of including false positives.

The results of our study further indicate that despite increased scrutiny by regulators and policymakers, the popular social media platform Instagram continues to act as a conduit for opioid, controlled substance, and illicit drug access, a direct violation of the RHA. Importantly, users have active conversations about selling and buying drugs, meaning that these social media posts act as digital marketplaces for drug dealing. Further action is needed to protect the public but needs to be carried out through meaningful collaboration and coordination involving partnership between technology companies, researchers, regulators, law enforcement, and impacted user communities to ensure that the opioid epidemic is not exacerbated by the digital risk environment.

Acknowledgments

TM would like to thank the US FDA for their invitation to present research associated with this study at the 2018 Online Opioid Summit and the FDA's leadership in this area.

Authors' Contributions

All authors jointly collected the data, designed the study, conducted the data analyses, and wrote the manuscript. All authors contributed to the formulation, drafting, completion, and approval of the final manuscript.

Conflicts of Interest

TM has received funding a pilot grant to research social media, illicit online pharmacies, and antibiotics from the Alliance for Safe Online Pharmacies, a 501(c)(4) social welfare organization engaged on the issue of illicit online pharmacies. TM is also the cofounder of S-3 Research, LLC, a company that received funding through the National Institutes of Health, National Institute on Drug Abuse—Start a SUD Startup Challenge award, and is also supported by an Accelerating Innovations to Market award provided by the University of California, San Diego—Jacobs School of Engineering Institute for the Global Entrepreneur. S-3 Research, LLC, has not received any nonpublic or industry funds related to its activity. TM and JL also received a pilot research grant from Google LLC to detect and characterize illegal drug dealing on YouTube. TM has previously received travel support from the World Health Organization and the US FDA to attend and participate in technical meetings regarding pharmaceutical supply chain and online drug safety. This includes travel support to attend and present at the FDA Online Opioid Summit as referenced in this paper. He has also acted as an expert consultant for the US Department of Justice on pharmaceutical diversion matters. None of these organizations had any involvement or input into the conception, design, collection, planning, conduct, analysis, interpretation, writing, and discussion of this work. Other coauthors have no conflicts or financial interests to disclose.

Multimedia Appendix 1

List of hashtag terms used.

[PDF File (Adobe PDF File), 114KB - jmir_v21i6e13803_app1.pdf]

References

1. Food and Drug Administration. 2018. FDA Invites Internet Stakeholders to Opioid Summit URL: https://www.fda.gov/drugs/newsevents/ucm608633.htm [accessed 2019-02-19] [WebCite Cache ID 76JE4mYcD]
2. Mackey TK. Opioids and the internet: convergence of technology and policy to address the illicit online sales of opioids. Health Serv Insights 2018;11:1178632918800995 [FREE Full text] [doi: 10.1177/1178632918800995] [Medline: 30245569]
3. House of Representatives, Committee on Energy and Commerce. Centers for Disease Control and Prevention. 2019 Feb 19. Wide-ranging online data for epidemiologic research (WONDER) URL: https://wonder.cdc.gov/ [accessed 2019-02-18] [WebCite Cache ID 76JWOwzKf]
4. Mackey TK, Liang BA, Strathdee SA. Digital social media, youth, and nonmedical use of prescription drugs: the need for reform. J Med Internet Res 2013 Jul 26;15(7):e143 [FREE Full text] [doi: 10.2196/jmir.2464] [Medline: 23892156]
5. Congress.gov.: Congress.gov; 2008. Ryan Haight Online Pharmacy Consumer Protection Act of 2008 URL: https://www.congress.gov/110/plaws/publ425/PLAW-110publ425.pdf [accessed 2019-02-19] [WebCite Cache ID 76JEKQcEd]
6. Mackey T, Kalyanam J, Klugman J, Kuzmenko E, Gupta R. Solution to detect, classify, and report illicit online marketing and sales of controlled substances via Twitter: using machine learning and web forensics to combat digital opioid access. J Med Internet Res 2018 Dec 27;20(4):e10029 [FREE Full text] [doi: 10.2196/10029] [Medline: 29613851]
7. Mackey TK, Kalyanam J, Katsuki T, Lanckriet G. Twitter-based detection of illegal online sale of prescription opioid. Am J Public Health 2017 Dec;107(12):1910-1915. [doi: 10.2105/AJPH.2017.303994] [Medline: 29048960]

XSL•FO
RenderX

8. Dasgupta N, Freifeld C, Brownstein JS, Menone CM, Surratt HL, Poppish L, et al. Crowdsourcing black market prices for prescription opioids. J Med Internet Res 2013;15(8):e178 [FREE Full text] [doi: 10.2196/jmir.2810] [Medline: 23956042]
9. Pergolizzi JJ, LeQuang JB, Taylor RJ, Raffa R, NEMA Research Group. The "Darknet": the new street for street drugs. J Clin Pharm Ther 2017 Dec;42(6):790-792. [doi: 10.1111/jcpt.12628] [Medline: 28921578]
10. Broséus J, Rhumorbarbe D, Mireault C, Ouellette V, Crispino F, Décary-Hétu D. Studying illicit drug trafficking on Darknet markets: structure and organisation from a Canadian perspective. Forensic Sci Int 2016 Jul;264:7-14. [doi: 10.1016/j.forsciint.2016.02.045] [Medline: 26978791]
11. Drug Enforcement Administration (DEA). Drug Enforcement Administration. Counterfeit Prescription Pills Containing Fentanyls: A Global Threat URL: https://www.dea.gov/sites/default/files/docs/Counterfeit%2520Prescription%2520Pills.pdf [accessed 2019-02-19] [WebCite Cache ID 76JXRyq3H]
12. Lytvynenko J. Buzzfeednews. 2018 Jun 1. Social Networks Are Losing A Deadly Battle With Illegal Online Pharmacies URL: https://www.buzzfeednews.com/article/janelytvynenko/social-networks-are-losing-a-deadly-battle-with-illegal [accessed 2019-02-19] [WebCite Cache ID 76JTBFyAp]
13. Smith A, Anderson M. Pew Internet. 2018 Mar 1. Social Media Use in 2018 URL: http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/ [accessed 2019-02-19] [WebCite Cache ID 76JTO6qoa]
14. United States Government Accountability Office. 2013 Jul. Internet Pharmacies Federal Agencies and States Face Challenges Combating Rogue Sites, Particularly Those Abroad URL: http://www.gao.gov/assets/660/655751.pdf [accessed 2019-02-19] [WebCite Cache ID 76JTgOf4I]
15. Orizio G, Merla A, Schulz PJ, Gelatti U. Quality of online pharmacies and websites selling prescription drugs: a systematic review. J Med Internet Res 2011 Sep 30;13(3):e74 [FREE Full text] [doi: 10.2196/jmir.1795] [Medline: 21965220]
16. Constine J. Techcrunch. 2018. Facebook cracks down on opioid dealers after years of neglect URL: https://techcrunch.com/2018/08/16/facebook-opioid-searches/ [accessed 2019-02-19] [WebCite Cache ID 76JTu0QWS]
17. Yang X, Luo J. Tracking illicit drug dealing and abuse on Instagram using multimodal analysis. ACM Trans Intell Syst Technol 2017 Jul 17;8(4):1-15. [doi: 10.1145/3011871]
18. Lapowsky I. WIRED. Tech companies deflect blame for opioid crisis ahead of FDA Summit URL: https://www.wired.com/story/tech-companies-deflect-blame-for-opioid-crisis-fda-summit/ [accessed 2019-02-19] [WebCite Cache ID 76JU9rYup]
19. Dwoskin E. The Washington Post. 2018 Sep 25. Instagram has a drug problem. Its algorithms make it worse URL: https://tinyurl.com/y65rvu6w [accessed 2019-02-19] [WebCite Cache ID 76JUMWJbR]
20. Tiku N. WIRED. 2018 Apr 6. One woman got facebook to police opioid sales on instagram URL: https://www.wired.com/story/one-woman-got-facebook-to-police-opioid-sales-on-instagram/ [accessed 2019-02-19] [WebCite Cache ID 76JUWUVmi]
21. Mackey TK, Kalyanam J. Detection of illicit online sales of fentanyls via Twitter. F1000Res 2017;6:1937 [FREE Full text] [doi: 10.12688/f1000research.12914.1] [Medline: 29259769]
22. Hanson C, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. J Med Internet Res 2013 Apr 17;15(4):e62 [FREE Full text] [doi: 10.2196/jmir.2503] [Medline: 23594933]
23. Hu H, Phan N, Geller J, Iezzi S, Vo H, Dou D. arXiv. 2019 Apr 3. An Ensemble Deep Learning Model for Drug Abuse Detection in Sparse Twitter-Sphere URL: https://arxiv.org/pdf/1904.02062.pdf [accessed 2019-06-05] [WebCite Cache ID 78ucr60Ib]
24. Cavazos-Rehg PA, Krauss MJ, Sowles SJ, Bierut LJ. Marijuana-related posts on Instagram. Prev Sci 2016 Aug;17(6):710-720. [doi: 10.1007/s11121-016-0669-9] [Medline: 27262456]
25. Cavazos-Rehg P, Krauss M, Fisher S, Salyer P, Grucza R, Bierut L. Twitter chatter about marijuana. J Adolesc Health 2015 Feb;56(2):139-145 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.10.270] [Medline: 25620299]
26. Katsuki T, Mackey TK, Cuomo R. Establishing a link between prescription drug abuse and illicit online pharmacies: analysis of Twitter data. J Med Internet Res 2015 Dec 16;17(12):e280 [FREE Full text] [doi: 10.2196/jmir.5144] [Medline: 26677966]
27. Hanson C, Cannon B, Burton S, Giraud-Carrier C. An exploration of social circles and prescription drug abuse through Twitter. J Med Internet Res 2013 Sep 6;15(9):e189 [FREE Full text] [doi: 10.2196/jmir.2741] [Medline: 24014109]
28. Zhou Y, Sani N, Lee C, Luo J. arXiv. 2016. Understanding Illicit Drug Use Behaviors by Mining Social Media URL: https://arxiv.org/pdf/1604.07096.pdf [accessed 2019-06-05] [WebCite Cache ID 78ueLIogr]
29. Huynh T, He Y, Willis A, Rüger S. Adverse Drug Reaction Classification With Deep Neural Networks. 2016 Presented at: COLING 2016; December 13-16, 2016; Osaka, Japan p. 877-887 URL: https://coling2016.anlp.jp/doc/main.pdf
30. Wu C, Wu F, Liu J, Wu S, Huang Y, Xie X. Detecting Tweets Mentioning Drug Name and Adverse Drug Reaction with Hierarchical Tweet Representation and Multi-Head Self-Attention. 2018 Presented at: 2018 EMNLP Workshop SMM4H: The 3rd Social Media Mining for Health Applications Workshop & Shared Task. 2018; October 31–November 1, 2018; Brussels, Belgium p. 34-37 URL: https://www.aclweb.org/anthology/W18-5909
31. Constine J. TechCrunch. 2018. Dealers remain on Instagram as it pushes drug searchers to treatment URL: https://techcrunch.com/2018/09/12/instagram-drugs/ [accessed 2019-06-05] [WebCite Cache ID 78ufyAbLT]
32. Quinlan J. Induction of decision trees. Mach Learn 1986 Mar;1(1):81-106. [doi: 10.1007/BF00116251]
33. Breiman L. Random forests. Mach Learn 2001 Oct;45(1):5-32 [FREE Full text]

XSL•FO
RenderX

34. Wang L. Support Vector Machines: Theory And Applications (Studies In Fuzziness And Soft Computing). Berlin: Springer; 2005.
35. Joachims T. Text categorization with Support Vector Machines: Learning with many relevant features. In: Machine Learning: ECML-98. Berlin, Heidelberg: Springer-Verlag New York, LLC; 1998:137-142.
36. Màrquez L, Rodríguez H. Part-of-speech tagging using decision trees. In: Machine Learning: ECML-98. Berlin, Heidelberg: Springer-Verlag New York, LLC; 1998:25-36.
37. Xu B, Guo X, Ye Y, Cheng J. An improved random forest classifier for text categorization. J Comput 2012 Dec 1;7(12):2913-2920 [FREE Full text] [doi: 10.4304/jcp.7.12.2913-2920]
38. Mikolov T, Karafiát M, Burget L, Černocký J, Khudanpur S. Recurrent Neural Network Based Language Model. 2010 Sep Presented at: INTERSPEECH 2010; September 26-30, 2010; Makuhari.
39. Sundermeyer M, Schlüter R, Ney H. LSTM Neural Networks for Language Modeling. 2012 Sep Presented at: INTERSPEECH 2012; September 9-13, 2012; Portland.
40. Chen Z, Watanabe S, Erdogan H, Hershey J. Speech Enhancement and Recognition Using Multi-Task Learning of Long Short-Term Memory Recurrent Neural Networks. 2015 Presented at: INTERSPEECH 2015; September 6-10, 2015; Dresden.
41. Weninger F, Erdogan H, Watanabe S, Vincent E, Roux J, Hershey J, et al. Speech enhancement with LSTM recurrent neural networksits application to noise-robust ASR. In: Latent Variable Analysis and Signal Separation. Cham: Springer, Cham; 2015:91-99.
42. Lin M, Chen Q, Yan S. arXiv. 2013. Network in network URL: https://arxiv.org/pdf/1312.4400.pdf [accessed 2019-06-05] [WebCite Cache ID 78uliSakV]
43. de Boer P, Kroese DP, Mannor S, Rubinstein RY. A tutorial on the cross-entropy method. Ann Oper Res 2005 Feb;134(1):19-67. [doi: 10.1007/s10479-005-5724-z]
44. Vogl TP, Mangis JK, Rigler AK, Zink WT, Alkon DI. Accelerating the convergence of the back-propagation method. Biol Cybern 1988 Sep 1;59(4-5):257-263. [doi: 10.1007/BF00332914]
45. Kingma D, Ba J. arXiv. 2014 Dec. Adam: A method for stochastic optimization URL: https://arxiv.org/pdf/1412.6980.pdf [accessed 2019-06-05] [WebCite Cache ID 78um3S6jI]
46. Portman R, Carper C. United State Senate. Combatting the opioid crisisxploiting Vulnerabilities in international mail URL: https://tinyurl.com/y5ky63b3 [accessed 2019-02-19] [WebCite Cache ID 76JUsthpf]
47. Zamora. House of Representatives, Committee on Energy and Commerce. Washington, D.C; 2018 Apr 11. Facebook: Transparency and use of consumer DATA URL: https://docs.house.gov/meetings/IF/IF00/20180411/108090/HHRG-115-IF00-Transcript-20180411.pdf [accessed 2019-02-19] [WebCite Cache ID 76JW9H39W]
48. O'Brien S. CNN Business. 2018 Apr 11. Mark Zuckerberg takes heat over illegal opioid sales on Facebook URL: https://tinyurl.com/y5h4bx6u [accessed 2019-02-17] [WebCite Cache ID 76JVEI7Cp]

Abbreviations

API: application programming interface
AUC: area under the curve
FCN: fully connected network
FDA: Food and Drug Administration
JSON: JavaScript Object Notation
LSD: lysergic acid diethylamide
LSTM: long short-term memory
MDMA: 3,4-methylenedioxy-methamphetamine
RF: random forest
RHA: Ryan Haight Online Pharmacy Consumer Protection Act, 2008
SVM: support vector machine

Edited by G Eysenbach; submitted 22.02.19; peer-reviewed by A Nara, J Kim, N Okazaki, N Miyoshi, Z Huang, JT Pollettini; comments to author 20.03.19; revised version received 12.05.19; accepted 27.05.19; published 15.06.19.

Please cite as:
Li J, Xu Q, Shah N, Mackey TK
A Machine Learning Approach for the Detection and Characterization of Illicit Drug Dealers on Instagram: Model Evaluation Study
J Med Internet Res 2019;21(6):e13803
URL: http://www.jmir.org/2019/6/e13803/
doi:10.2196/13803
PMID:31199298

XSL•FO
RenderX

©Jiawei Li, Qing Xu, Neal Shah, Tim K Mackey. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Early Detection of Depression: Social Network Analysis and Random Forest Techniques

Fidel Cacheda[1,2], PhD; Diego Fernandez[1,2], PhD; Francisco J Novoa[1,2], PhD; Victor Carneiro[1,2], PhD

[1]Department of Computer Science, Faculty of Computer Science, University of A Coruna, A Coruna, Spain

[2]Center for Information and Communications Technology Research, University of A Coruna, A Coruna, Spain

Corresponding Author:
Diego Fernandez, PhD
Department of Computer Science
Faculty of Computer Science
University of A Coruna
Campus de Elvina
A Coruna, 15071
Spain
Phone: 34 881011213
Email: diego.fernandez@udc.es

Abstract

Background: Major depressive disorder (MDD) or depression is among the most prevalent psychiatric disorders, affecting more than 300 million people globally. Early detection is critical for rapid intervention, which can potentially reduce the escalation of the disorder.

Objective: This study used data from social media networks to explore various methods of early detection of MDDs based on machine learning. We performed a thorough analysis of the dataset to characterize the subjects' behavior based on different aspects of their writings: textual spreading, time gap, and time span.

Methods: We proposed 2 different approaches based on machine learning singleton and dual. The former uses 1 random forest (RF) classifier with 2 threshold functions, whereas the latter uses 2 independent RF classifiers, one to detect depressed subjects and another to identify nondepressed individuals. In both cases, features are defined from textual, semantic, and writing similarities.

Results: The evaluation follows a time-aware approach that rewards early detections and penalizes late detections. The results show how a dual model performs significantly better than the singleton model and is able to improve current state-of-the-art detection models by more than 10%.

Conclusions: Given the results, we consider that this study can help in the development of new solutions to deal with the early detection of depression on social networks.

(J Med Internet Res 2019;21(6):e12554) doi:10.2196/12554

KEYWORDS

depression; major depressive disorder; social media; artificial intelligence; machine learning

Introduction

Background

Major depressive disorder (MDD), also known simply as depression, is among the most prevalent psychiatric disorders globally [1,2]. As described in the World Health Organization's Comprehensive Mental Health Action Plan 2013-2020 [3], depression alone affects more than 300 million people worldwide and is one of the largest single causes of disability worldwide, particularly for women. Depression currently accounts for 4.3% of the global burden of disease, and it is expected to be the leading cause of disease burden in high-income countries by 2030 [4].

The Institute of Medicine Committee on the Prevention of Mental Disorders identified depression as the most preventable disorder [5], and several studies have demonstrated that early recognition and treatment of depression can improve the negative impacts of the disorder [6-8]. Therefore, it is vital to provide an early identification of subjects suffering from depression to intervene as soon as possible and minimize the impact on public health by potentially reducing the escalation of the disease. However, provisions and services for the early

XSL•FO
RenderX

detection and treatment of depression and other mental health disorders remain limited. Although there are also some validated laboratory tests to diagnose depression, such as Beck Depression Inventory - II, Center for Epidemiologic Studies Depression Scale (CES - D), Geriatric Depression Scale, Hospital Anxiety and Depression Scale, Patient Health Questionnaire - 9 [9,10], and Hamilton Rating Scale for Depression [11] most diagnoses are formed on the basis of self- or family reports.

In this context, the relation between language and clinical disorders has been analyzed for years [12,13]. Taking this into account, new work has appeared to predict and study depression [14,15]. In particular, researchers are increasingly examining the potential of social media networks as tools to predict depression and detect its symptoms as manifested in user comments and related activities. Social networks such as Twitter, Inc, Facebook, Inc, and Reddit, Inc have become part of our daily lives as media through which to share our thoughts, feelings, and overall emotional status. As such, these platforms have become valuable data banks for marketers and researchers, who can analyze user metrics, shared content, and related information to identify preferences and tastes as well as other attitudes and behaviors [15,16]. In fact, social networks have proved to be used by patients to interact with peers because of their support and ability to understand someone's experience, while maintaining a comfortable emotional distance [17]. For example, Reddit, Inc is an open-source platform where community members can submit content and vote on submissions. Content entries are organized by areas of interest (denoted as subreddits), with a large history of previous submissions covering several years. This social network is particularly interesting for our study, as it contains substantive content about different medical conditions, including MDD.

This study uses publicly available data from Reddit, Inc to examine the effectiveness of different methods that can provide an early detection of MDDs based on artificial intelligence. As detailed in the next sections, we mainly focus on 2 different methods, both of which are based on machine learning algorithms that use textual and semantic similarity features along with writing features (WFs) to predict a subject's depression condition. The first technique follows a simpler proposal using a single machine learning algorithm, whereas the second model follows a dual approach that uses 2 machine learning algorithms: the first one is trained to predict depression cases, whereas the second one is trained to predict nondepression cases. We conducted a thorough evaluation of each model following a time-aware approach that rewards early detections and considers late detections as false negatives. Our results show that the dual model can improve state-of-the-art detection models up to 10%. Furthermore, our methods were implemented using freely available tools, thus facilitating the reproduction of our research work [18].

The aim of this study was to explore the use of machine learning for an early detection of MDD using WFs from social network content to improve state-of-the-art methods, which can lead to the development of early detection technologies that could help in the identification of subjects suffering from depression. The main contributions of our study can be summarized as follows:

- We provide a detailed analysis on publicly available data from social networks to characterize the subjects' behavior based on different aspects of their writings: textual spreading, time gap, and time span.
- We propose 2 different machine learning methods, named singleton and dual, that use textual, semantic, and WFs derived from subjects' social networks behavior to predict his depression condition.
- We follow a time-aware evaluation that strictly penalizes late depression detections. Our results show that the dual model is able to improve upon state-of-the-art methods.

The structure of the paper is as follows. First, we examine related studies with regard to early detection of depression with a particular focus on techniques that use information extracted from social networks. Then, we provide a detailed data analysis of the social network content for MDD detection and we describe our proposed model for the early detection of depression. After the methods, we present the results and performance improvements obtained over the state-of-the-art baselines. Finally, we summarize our conclusions and future studies in this line of research.

Related Studies

Several previous studies have highlighted the importance of early detection in improving outcomes related to MDD [6-8]. Halfin's study [6] demonstrated that the early detection, intervention, and appropriate treatment can promote remission and reduce the emotional and financial burdens of this disease, and Picardi et al [7] observed significant improvements in depressive symptoms and quality of life among subjects who had undergone early screening. Rost et al [8] found that early intervention for depression can improve employee productivity and reduce absenteeism.

Over the past decade, social networks have increasingly become a focus of research efforts to identify and characterize the incidence of various disorders. For example, Prieto et al [19] proposed a method to use Twitter, Inc to automatically measure the incidence of a set of health conditions. Chunara et al [20] analyzed cholera-related tweets published during the first 100 days of the 2010 Haitian cholera outbreak, and Chew and Eysenbach [21] used sentiment analysis on 2 million tweets to propose a complementary infoveillance approach. Aladağ et al [22] have studied posts looking for regular language patterns to prevent potential suicide attempts. Even Rice et al [23] have demonstrated that the development of cost-effective, acceptable, and population-focused interventions is critical in depression. A number of online interventions (both prevention and acute phase) have been tested in young people with promising results.

Diverse studies have explored the potential of social media networks to predict and detect mental health disorders [24-28]. For example, De Choudhury et al [27] developed a statistical methodology to derive distinct markers of shifts to suicidal ideation from Reddit, Inc user data for modeling in a prediction framework, and Birnbaum et al [25] proposed a method that used machine learning in combination with clinical appraisals as a means of identifying social media markers of schizophrenia.

XSL•FO
RenderX

Other studies have focused specifically on depression. Ziemer and Korkmaz's [29] comparison of human versus automated text analyses of psychological and physical disorders found human ratings of depression to be more accurate than machine-based methods; however, other studies have yielded promising, albeit limited, results using sophisticated technological applications in detecting and measuring the disorder. Nadeem's *bag of words* analysis of Twitter, Inc messages [30] examined the frequency of use of *my* and *me* as a marker for depression, whereas De Choudhury et al [15] leveraged social activity, emotion, and language signals manifested on Twitter, Inc to introduce a social media depression index. Similarly, a task organized at the Computational Linguistics and Clinical Psychology Workshop 2015 to detect depression and other mental health disorders among subjects using Twitter, Inc posts achieved promising results using topic modeling and rule-based methods [31-33].

Fewer studies have focused on early detection of depression. Ophir et al [34] examined signals of depression among adolescent Facebook, Inc users with the aim of ultimately applying their coding scheme to early detection methods, although no methods are proposed by the authors. De Choudhury et al [15] achieved 70% accuracy in an experiment that compared scores found on the Center for Epidemiologic Studies Depression Scale [35] and BDI [36] with Twitter, Inc users' engagement patterns and linguistic markers preceding a recent episode of depression to devise a tool for predicting and measuring MDD in individuals. This study identified several distinctive features of posting activity associated with the onset of depression, such as diurnal cycles, more negative emotions, less social interaction, more self-focus, and more mentions of depression-related terms. However, as with most other research that attempts to predict depression, the analysis was dependent on self-reported cases, and to date, approaches aiming to identify individuals who are as yet unaware of their depression diagnosis remain rare [28]. Moreover, in this study, the authors did not perform an early detection evaluation.

Our study is directly related to the Conference and Labs for the Evaluation Forum workshop on early risk prediction on the internet (eRisk) 2017 [37], during which the authors proposed a task on the early detection of depression with a time-aware methodology and using effectiveness metrics. In general, participants based their approaches on lexical, linguistic, semantic, or statistical features, among others. We followed the workshop methodology [13,37] and used the best performing methods as baselines [38-39]. Trotzek et al [38] based their model on linguistic metainformation extracted from the subjects' writings and developed a classifier using recurrent neural networks, whereas Villegas et al [39] explicitly modeled partial information from the semantic representation of documents using learning algorithms such as random forest (RF) or naive Bayes. Our study follows the same evaluation methodology as these studies, but it diverges from them in being a dual-model proposal, as well as in terms of the specific WFs analyzed.

Methods

Data Analysis

Our input comprised a set of posts and comments from a social network, specifically gathered for eRisk 2017 [13]. Data were extracted from Reddit, Inc using the Reddit, Inc's application program interface (API), and the resulting dataset consists of a collection of tuples of the form (id, writing), such that *id* is a unique identifier for each social network user and *writing* represents a writing instance in the social network. At the same time, each writing was described as a tuple of the form (title, date, info, and text), whereby *title* indicates the title of the post or comment, *date* denotes the date and time when the writing was performed, *info* identifies the social network (in this case, only Reddit, Inc is considered), and *text* comprises the actual post or comment provided by the user. The *title* value of a comment is empty, as, in this case, the user is replying to a previous post (whose title is already defined).

Depressed users are identified by searching in the depression subreddit for posts with specific self-reports of diagnosed depression. These reports must include a more or less specific date of diagnosis. However, the errors committed in these dates are not going to interfere with the experiments because we aim at detecting if a user has been depressed or not, regardless of the concrete date of diagnosis. Moreover, a strict manual review was performed to verify that posts were genuine.

Then, a control group was created by randomly selecting a large set of redditors, including some individuals who were active on the depression subreddit but had no depression diagnosed [13]. It is important to remark that collaborating in the depression subreddit does not imply to be depressed. For instance, people trying to help others may participate in this subreddit.

The controls have not been checked for other diseases, and it is assumed that they are not depressed because they have not manifested their depression in their writings, the unique evidence used from Reddit, Inc. In fact, writings for control and depressed users are gathered from all the subreddits where the users had written, without paying attention to the concrete issues. Only users with at least 10 submissions have been considered.

The dataset has been formed starting from those writings where users claimed that they were depressed [13]. From there, a period of about a year has been considered for each user. The intervals can differ because the maximum amount of submissions that can be retrieved per redditor is 2000 (Reddit, Inc's API limit).

As shown in Table 1, the dataset includes a total of 887 subjects, of which 135 have been diagnosed with depression, and encompasses more than 500,000 different posts and comments, with an average of nearly 600 posts per subject. In addition, other descriptive statistics are shown to demonstrate the differences between control and depressed users. On the basis of these data, we focused on estimating the likelihood that a particular subject could be considered depressed given his particular social network posts.

XSL•FO
RenderX

Table 1. Analysis of dataset statistics.

Features	Depressed	Control	Total
Subjects, n	135	752	887
Posts, n	49,557	481,837	531,394
Number of submissions per subject			
Average	367.1	640.7	599.1
Median (range)	154 (10-1832)	375 (10-2000)	321 (10-2000)
Interquartile range	562	1039.5	1006
Average words per submission	27.3	21.9	22.4
Period of time per subject (days)			
Average	586.42	625.02	619.15
Median (range)	520.95 (0.60-2249.48)	477.12 (0.26-3067.16)	484.88 (0.26-3067.16)
Interquartile range	786.88	753.19	756.83

Subject Behavior

To characterize the subject's behavior on the dataset, we performed a detailed analysis of the main characteristics that might have an impact on the early detection of depression. We concentrated on variables that could be easily measured directly from the writings and in which we expected to capture certain differences in behavior between both types of subjects.

Textual Spreading

We began our analysis by characterizing the textual spreading of the writings produced by the subjects by measuring the number of words used in each of the writings. Figure 1 shows the words used in the post titles, both for depressed and nondepressed individuals. In particular, the number of titles with zero words (that is, comments to previous posts) is significantly higher among depressed users. That can be explained by considering how Reddit, Inc users can publish new writings: they can either publish a new reddit, for which it is mandatory to add a title; or they can comment on an already existing reddit. Thus, these results led us to conclude that depressed users have a higher tendency to reply to existing issues rather than publish new ones.

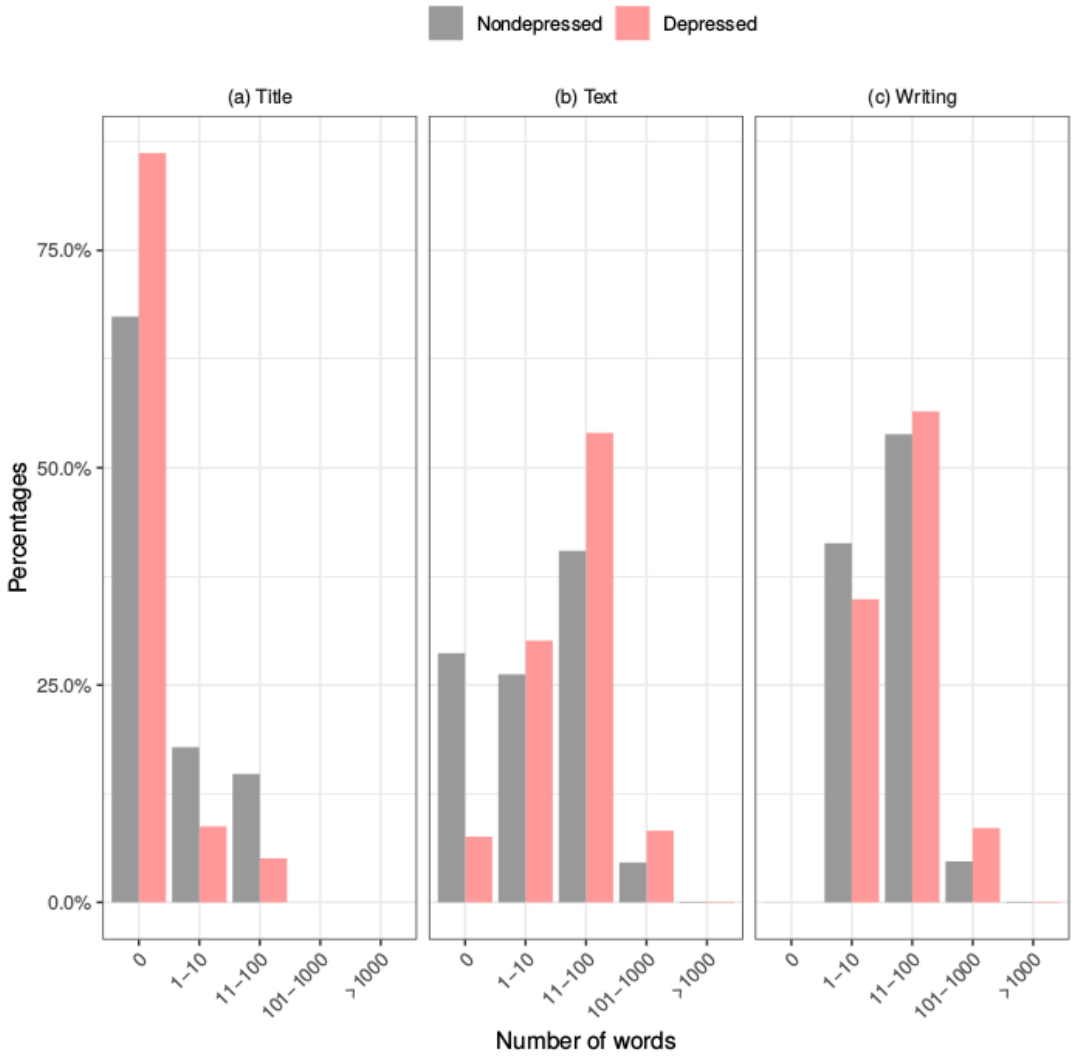
Conversely, analyzing the second plot in Figure 1, we can observe how the nondepressed users tended to send many more writings with zero words in their content description, whereas depressed subjects tended to elaborate more on their writings. In fact, the percentage of posts using between 11 and 100 words is nearly 14 points higher for depressed subjects, and it nearly doubles the percentage for even larger posts (more than 100 words). To better understand this analysis, it is important to note that there are 2 kinds of new submissions in Reddit, Inc: text submissions, whereby a user can add a text description to his title; and link submissions, in which text descriptions cannot be added, thus producing zero words in the text field.

The third plot in Figure 1 demonstrates that the total textual spreading of writings is similar for both depressed and nondepressed subjects. Although there are clear differences between the ways that depressed and nondepressed users submitted their writings, the differences in the titles are compensated for by the differences in the text, which results in similar distributions taking into account the total number of words. In any case, it is noticeable that the depressed individuals tended to elaborate their writings more and use more overall words than those who were not depressed.

These results have been checked by conducting different hypothesis contrasts. First, we employed 3 *F* tests studying the equality of variances for the number of words in title, text, and writing, considering control and depressed users. The results indicate that variances are different for title ($P<.001$) and text ($P<.001$) but equal considering the whole writing ($P=0.62$). Regarding the means, the Student *t* test computation resulted in accepting the alternative hypothesis, so the means are not equal. The *P* value is <.001 for these 3 contrasts.

XSL•FO
RenderX

Figure 1. Relative percentage for number of words used on title (a), text (b), and both fields (c) for depressed and nondepressed individuals.



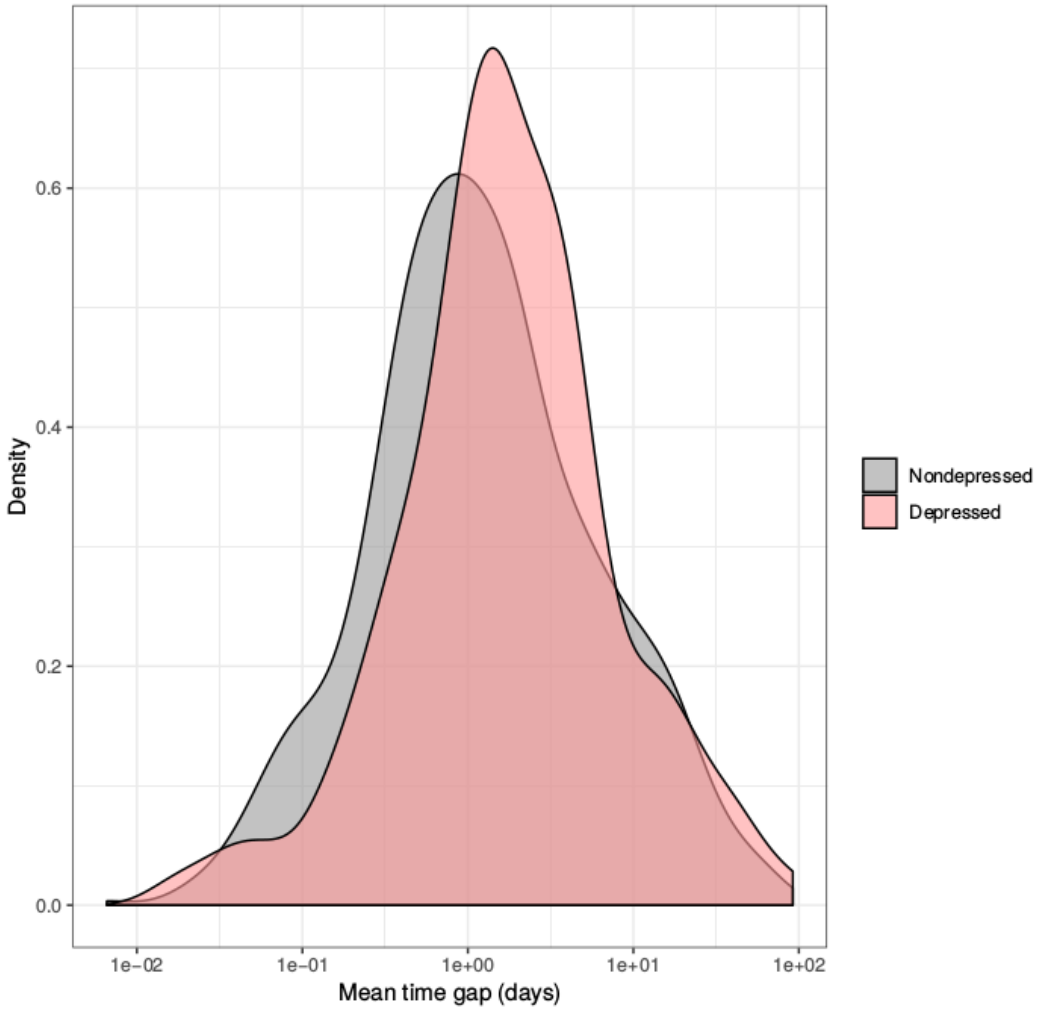
Time Gap

Next, we focused on the time gap between 2 consecutive writings. Figure 2 displays a density plot for the time gap between writings for depressed and nondepressed individuals. In Figure 1, we can observe a higher mean among the depressed subjects, taking more time between 2 consecutive writings. In fact, the average time spent for a depressed subject between 2 writings is 5 days (5.076), whereas nondepressed writers will post again 1 day faster (4.037). In addition, the differences in the SD are significant, which is about 8 days (8.330) for nondepressed subjects but rises to 11 days (11.048) for depressed subjects. This result suggests that depressed subjects exhibit higher variability in their publication routine on the social network.

Starting from the logarithmic values of the time gap, the equality of variances was tested using an *F* test contrast. The resultant *P* value was .52, so variances are equal. In addition, the means were tested for equality between both subject types using 2-sided *t* test with significance level alpha=.05, showing that means are different (*P*=.02), which confirmed significant differences between these values among depressed and nondepressed subjects.

XSL•FO
RenderX

Figure 2. Average time gaps distribution between writings for depressed and nondepressed subjects.



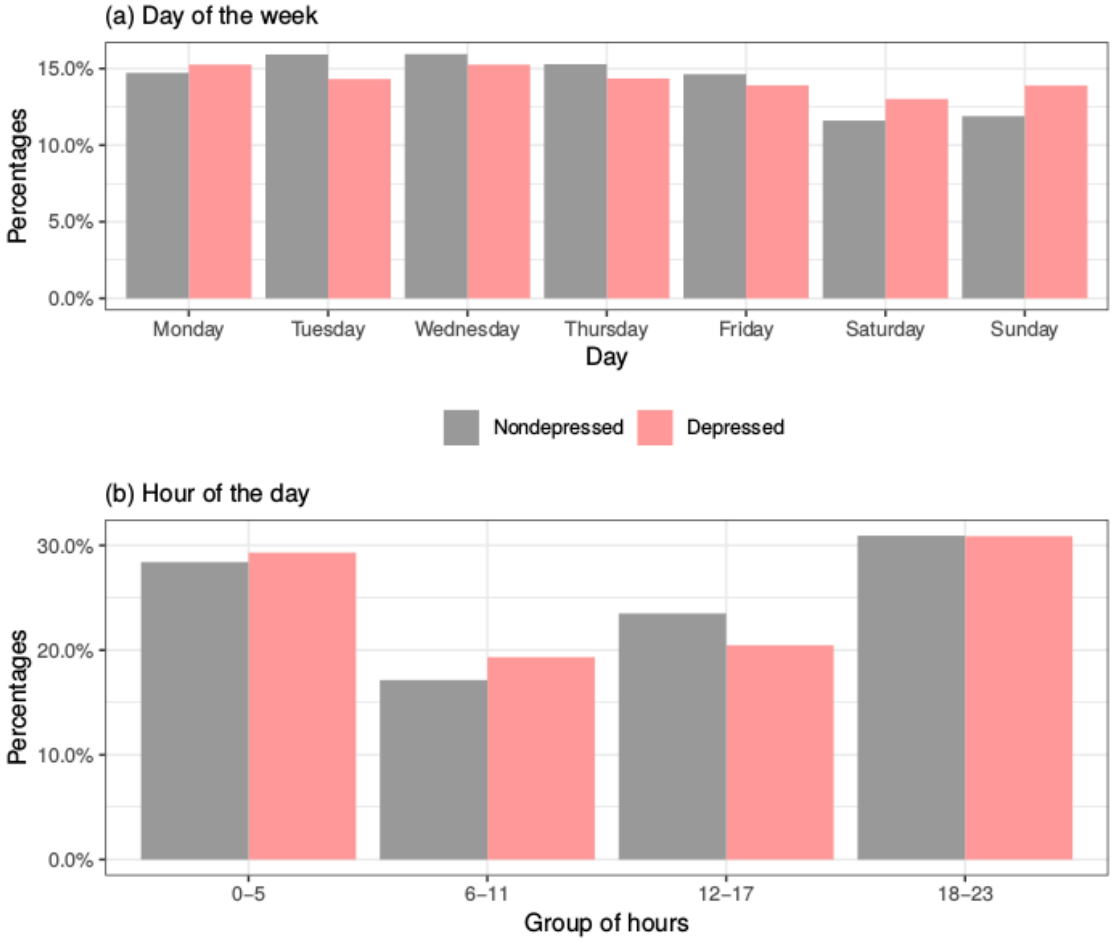
Time Span

We also explored the time span of the different writings in terms of the days and times of day when they were produced. The classification of writings according to the day of the week is described in the first plot in Figure 3. The main difference between both types is that nondepressed subjects tended to publish less during weekends than depressed individuals, whereas this tendency was inverted during weekdays, except on Mondays. In general terms, the publication rate is more homogeneous for depressed individuals, despite a small reduction at weekends. The nondepressed subjects exhibited a publication peek on Wednesdays, followed by a gradual reduction that reaches its lowest point during the weekend.

Finally, the second plot in Figure 3 shows how depressed subjects tended to send more posts and comments than nondepressed users over the hours from midnight to midday, whereas the latter published more in the afternoon. The main differences appear 6 hours before midday, when depressed subjects were most active, and 6 hours after, when nondepressed subjects were most active. The same behavior was observed by Choudhury et al [15], arguing that online activity at night is a known characteristic of these individuals, which may be the reason behind this increase.

XSL•FO
RenderX

Figure 3. Time span bar plots according to the day of the week (a) and hour of the day (b) for depressed and nondepressed subjects.



Depression Prediction

The depression prediction problem presented in this study can be formalized as a binary classification problem using the presence or absence of depression diagnosis as a label. Accordingly, to address this machine learning problem, we resorted to a features-based approach and designed a collection of features that are expected to capture correlations between different aspects of the individual's writings and depression. We represented each training example by a feature vector: ϕ (id, writing) $\in R^F$, where F denotes the number of features, and then, we used this vector as input for the prediction function V. Using this approach enabled us to develop a large number of features, and we employed techniques suited for learning on a large scale, such as tree-based algorithms, to estimate relationships between those features and depression. We proposed 3 types of features: textual similarity, semantic similarity, and WFs.

Textual Similarity Features

Positive subjects refer to those diagnosed with MDD and vice versa for negative subjects. The main goal of these features is to estimate the degree of alignment of a subject's writings with those of positive or negative subjects, which enables the researcher to estimate the similarity between a given subject versus positive and negative subjects. We ignored word ordering and opted for a bag-of-words representation that considered 2 different measures extensively used in the literature: cosine similarity (an instantiation of a vector space model [VSM]) and Okapi Best Matching 25 (BM25, an instantiation of a probabilistic model). The former calculates the angle formed by 2 term-frequency vectors, whereas the latter tries to estimate the probability of relevance between a query and a document.

Each subject was represented as a document that included all his writings and was modeled as a collection of words: $d=\{w_1, \ldots, w_{l(d)}\}$, where $l(d)$ represents the number of terms in the text.

The cosine similarity between 2 subjects q and d is calculated as in equation a in Figure 4 following the study by Singhal [40], where $cnt(q_i, q)$ is the number of times that the term q_i appears in the document q and IDF(q_i) is the inverse document frequency for term q_i that is computed over a corpus C as specified in equation b in Figure 4. In this equation, $n\_docs\ s\ (C)$ represents the overall number of documents in C (equivalent to the number of subjects), whereas $n\ (w; C)$ is the number of documents that contain the term w.

XSL•FO
RenderX

Figure 4. Textual similarity measures. IDF: inverse document frequency; BM25: Okapi Best Matching 25.

$$(a)\ cos(q,d) = \frac{\sum_{i=1}^{l(q)}(IDF(q_i)\cdot cnt(q_i,q))\cdot(IDF(d_i)\cdot cnt(d_i,d))}{\sqrt{\sum_{i=1}^{l(q)}(IDF(q_i)\cdot cnt(q_i,q))^2}\cdot\sqrt{\sum_{i=1}^{l(d)}(IDF(d_i)\cdot cnt(d_i,d))^2}}$$

$$(b)\ IDF(w;C) = log\frac{n_{docs(C)}-n(w;C)+0.5}{n(w;C)+0.5}$$

$$(c)\ BM25(q,d) = \sum_{i=1}^{l(q)} IDF(q_i)\frac{cnt(q_i,d)\cdot(k_1+1)}{cnt(q_i,d)+k_1(1-b+b\cdot l(d)/\overline{l})}$$

The Okapi BM25 similarity between 2 subjects q and d was scored as equation c in Figure 4, following the study by Robertson and Zaragoza [41], where k_1 is a scaling factor for the term frequency, b is a scale factor on the document length, and l is the average number of terms in a document. In our setting, each subject was represented by the concatenation of all the textual information available for each writing (title, info, and text), and the inverse document frequency dictionary was computed over the overall collection of these documents. The textual similarity between two subjects might have different degrees of importance, and to address this effect, we also used the aggregation of cosine and BM25 scores, with each computed between the different parts of the textual information available.

We aggregated the scores obtained for each individual's writings by calculating the average value, SD, minimum, maximum, and median. This process was repeated for both positive and negative samples and, in all cases, the active subject was removed from the samples.

Semantic Similarity Features

We applied latent semantic analysis (LSA) as one of the best known VSMs to capture semantic relationships among documents. LSA explicitly learns semantic word vectors by applying singular value decomposition, which in turn projects the input word representation into a lower-dimensional space of dimensionality $k << V$, where semantically related words are closer than unrelated words.

In LSA, a document-term matrix M was constructed from a given text base of n documents containing m terms. This matrix of size m x n was then decomposed via a singular value decomposition into 3 matrices: the term vector matrix T; the document vector matrix D, and the diagonal matrix S:

$$M=TSD^T \quad (1)$$

These matrices were then reduced to the given number of dimensions k to result into truncated matrices T_k, S_k, and D_k, creating the latent semantic space [42], as specified in Figure 5.

Different dimensionality methods have been tested in the literature. To compute the k dimensionality, this study typically used the Kaiser criterion [43], which will take values higher than 1.0 and return the number of singular values accordingly. We also tested the share dimensionality, which finds the first position in descending order of the singular values where their proportional sum meets or exceeds a specific share, and the fraction dimensionality, which takes a specific fraction of the number of singular values [44]; however, no relevant differences were identified among the different methods.

Semantic similarity features between 2 subjects were computed as the Euclidean distance between the respective projections into the embedding space. As described in previous section (Textual Similarity Features) , each subject was represented as a document that aggregated all of his writings. In this case, all the available textual information was used to compute the singular values. LSA was applied both following a full-text approach and removing stop words and using Porter stemming [45]. Finally, we applied feature scaling to normalize the LSA scores computed following minimum-to-maximum normalization [46]:

$$x' = (x - min(x)) / (max(x) - min(x)) \quad (2)$$

In this equation, x is the original value and x' is the normalized value.

Writing Features

The collection of features was used to profile the characteristics of the subjects' writings on the basis of the findings from Data Analysis. As reviewed above, we defined 3 signals: textual spreading, time gap, and time span. Textual spreading measures the amount of textual information provided by the subject in his writings, and to address this feature, we introduced the following features:

- NWritings: The number of writings produced by the subject.
- AvgWords: The average number of words per writing. For each writing all the textual information available is considered.
- DevWords: SD for the number of words per writing.
- MinWords: Minimum number of words in the subject's writings.
- MaxWords: Maximum number of words in the subject's writings.
- MedWords: Median for the number of words in the subject's writing.

Figure 5. Latent semantic space.

$$M_k = \sum_{i=1}^{k} t_i \cdot s_i \cdot d_i^T$$

XSL•FO
RenderX

To measure the time elapsed between 2 consecutive writings, we aggregated the writings' time gap information for each subject. In this way, if a subject only had one writing in the time period considered, the time gap would be zero. Otherwise, the time gap would measure the number of seconds between 2 consecutive writings. A logarithmic transformation of the raw time gap values was also considered, resulting in the following 2 sets of features:

- TimeGap: The aggregated information for the time lapse between 2 consecutive writings. These values are represented as the average, SD, minimum, maximum, and median.
- LogTimeGap: For the logarithmic transformation of the time gap values. The same aggregation values are computed for each subject.

Another group of features was used to profile the moment when the writings were created by the subject. This information was expected to model differences in behavior among subjects diagnosed with depression versus those who had not been so diagnosed. The following time features were proposed:

- Day: Percentage of writings provided by the subject, for each day of the week.
- Weekday: Accumulative percentage for all writings created in a weekday.
- Weekend: Accumulative percentage for all writings posted during the weekend.
- Hour: The hours of the day are divided into 4 homogeneous classes (0:00-5:59, 6:00-11:59, 12:00-17:59, and 18:00-23:59) and the percentage of writings that fall into each class is calculated.

As a summary, textual and semantic features are computed and aggregated for each user in comparison with all other users (grouped as positive and negative), meanwhile WFs are independently calculated and aggregated for each individual with respect to his postings.

Models

We employed a readily available machine learning toolkit [47] to develop a learning model incorporating the features that were identified. We analyzed some standard machine learning algorithms (ie, C4.5, random tree, and RF) on this classification problem and selected RF [48] as the best performing model. An independent subsampling set was used to estimate the number of trees, and BM25's b and k_1 metaparameters.

The evaluation followed a time-aware methodology in which the writings were chronologically sorted and grouped into subsets. Each subset was evaluated independently, and the model was required to emit 1 of 3 possible decisions:

- Depression: The subject is considered to suffer from depression. This decision is final.
- Nondepression: The subject is considered not to suffer from depression. This decision is final.
- No decision: There is not enough evidence to produce a definitive decision and it is delayed.

As this is not a traditional binary classification problem because of the delay option available when processing the different subjects' writings, we proposed 2 different approaches: singleton and dual. The singleton model uses only 1 RF model, which is trained using the corresponding features, and a decision function is integrated to determine if enough evidence is available to proceed with a firm diagnosis or the decision must be delayed. The decision function was defined as $\delta(m, th_+(i), th_-(i))$, where m denotes the machine learning model used in the binary classification problem and $th_+(i)$ is a threshold function that sets a limit for a positive decision depending on the information chunk being processed (i), whereas $th_-(i)$ is a threshold function that sets a limit for a negative decision. Both threshold functions are not required to be the same, although they could be.

Different threshold functions were considered; however, the best performance was obtained with a decreasing step function. The steps of these threshold functions were tuned with a grid search over {0.95, 0.9, 0.85, 0.8, 0.75, 0.7, 0.65, 0.6, 0.55, 0.5} on the training set, and selected the best performing steps for experimentation. Finally, both threshold functions (positive and negative) are the same and follow the equation:

$$th(i) = 0.9_{XA[1,0.9]} + 0.8_{XA[0.9,0.8]} + 0.7_{XA[0.8,0.7]} + 0.6_{XA[0.7,0.6]} + 0.5_{XA[0.6,0.5]} \quad (3)$$

In previous equation, $XA(x)$ is the indicator function defined as 1 if x belongs to A, or 0 if x does not belong to A.

The main problem of the singleton approach is that it uses a binary classifier and to provide a final decision, both options (depressed or nondepressed) compete against each other and, therefore, require important support from the data features to surpass the threshold, thus causing a delay. Note that for 1 option (eg, depressed) to reach a probability of 0.9, the other option (eg, nondepressed) must be 0.1.

To overcome this matter, and inspired by the multiclass classifiers *one-versus-all* that train different binary models and select the most positive value [49], we propose the dual model that uses 2 RF models, each one trained with an independent set of features and, this way, both options do not compete but can be predicted independently. The first model (m_+) is trained to predict depression cases, whereas the second model (m_-) is trained to predict nondepression cases. For the dual model, a decision function of the form $\delta(m_+, m_-, th_w, th_+, th_-)$ was defined, where m_+ and m_- are the 2 learning models considered, thw denotes the number of threshold writings and th_+ and th_- are the threshold functions applied to m_+ and m_-, respectively. Both threshold functions are defined as constant functions of the form, where the value for th_+ is 0.9, and the value for th_- is 0.5.

The positive threshold function takes the upper step (0.9) from the positive threshold function of the singleton model, whereas the negative threshold function takes the lower step (0.5) of the negative threshold function of the singleton model. These thresholds were achieved following a grid search over the same values as the singleton model.

In the dual model, if the number of writings is below th_w, the first model is applied with decision threshold function th_+, so that if a positive probability is above the threshold, a depression

XSL•FO
RenderX

decision is emitted, otherwise the decision is delayed. If the number of writings is above the writings threshold, the second model is applied with decision threshold function th_{-}. In this case, if the nondepression probability is above the threshold the final decision is emitted and if otherwise, the decision is delayed. In this way, each classifier (m_{+} and m_{-}) operates with independent features and each one can, independently, reach the threshold and provide an earlier final decision.

Results

Dataset

Table 2 presents the main statistics for the dataset. A total of 892 subjects were considered, of whom approximately 15% had been diagnosed with MDD. All submissions were collected from Reddit, Inc for a period covering more than 1 year [13]. Subjects with less than 10 submissions were removed.

The following evaluation is based on a subject-based train-test split, as reported in Table 2, with an approximate percentage of 55% on the training set and 45% for testing.

The sequence of writings in the test set was chronologically sorted and the set was further divided into 10 subsets (or chunks), each of which contained 10% of the messages. These subsets were considered sequentially in such a manner that the first subset contained the oldest 10% messages, the second subset the second oldest 10%, and so forth. This test subset division was a particularly important element in the evaluation, as its main objective was to detect, as soon as possible, a depression case, which would represent an improvement over traditional evaluation, which identifies cases without regard for speed. This becomes patent in the performance measure described in the next section.

Performance Measure

Standard classification metrics such as precision, recall, or *F* measure do not take into account time, and therefore, we opted for early risk detection error (ERDE) [13]. This measure will consider both the correctness of the decision and the delay taken by the model to make the decision, where the delay is measured by the number of writings (posts or comments) seen before providing an answer.

Given a decision (d) taken by the system with a delay (k) and a ground truth (gt) for each subject, the ERDE measure is defined as equation a in Figure 6.

In that equation c_{fp} and c_{fn} are the costs associated with a false positive and false negative, respectively. In this study, following Losada and Crestani [13], c_{fn} was set to 1 and c_{fn} was set to the proportion of positive cases in the test dataset (ie, 0.1296). The correct detection of a negative does not have any repercussion (negative nor positive) in the performance of the system, independently of the moment when it is detected, as this is considered a nonrisk case that would not require an early intervention. In the case of a correct positive decision, the factor $lc_o(k)$ introduces a cost associated to the delay in detecting a true positive. As suggested by Losada and Crestani [13], $c_{tp}=cf_n$, as a late detection can have the same negative consequences as a false negative. For the $lc_o(k)$ factor, we use a monotonically increasing function of k as specified in equation b in Figure 6.

For each subject, the ERDE metric was computed, and a final score was obtained averaging all the ERDE values. As all cost weights are between 0 and 1, both included, then ERDE is also in the same range, and the quality of system performance increases as values approach 0. Following the evaluation procedure by Losada and Crestani [13], $ERDE_5$ and $ERDE_{50}$ measures were used for a comparison with the baselines, where 5 and 50 represent the subscript o for lc_o factor, that is, the number of writings processed from where ERDE increases more rapidly.

Baselines

Table 3 presents the main metrics for the baselines considered. The first 3 rows contain some naïve baseline methods, the middle rows show results for some Oracle methods, and the last 2 rows expose the best performing methods from eRisk 2017 [13]. For all methods, we present the $ERDE_5$ and $ERDE_{50}$ metrics as the performance measures used in the eRisk 2017 competition, as well as *F* measure, Precision, and Recall.

Three different naïve methods that do not require any specific features (textual, semantic, or writing) were considered. The random strategy emits a random decision for each subject. As the evaluation is divided into 10 chunks, this method produces a random and equally probable verdict (*depression*, *nondepression*, or *no decision*) for each subject at the end of each chunk. As soon as the system produces a diagnosis (*depression* or *nondepression*), later decisions are not taken into account. The naïve all-depressed method will emit a *depression* decision for all subjects for all chunks. As the first chunk provides a decision for all subjects, the actions in the following chunks do not have any repercussion in the system performance. In this case, the recall reached its maximum, as expected, although both ERDE metrics obtained modest results.

Table 2. Dataset statistics.

Features	Training		Test	
	Depressed	Control	Depressed	Control
Subjects, n	83	403	52	349
Posts, n	30,851	264,172	18,706	217,665
Average submissions per subject	371.7	655.5	359.7	623.7
Average words per submission	27.6	21.3	26.9	22.5

XSL•FO
RenderX

Figure 6. Early risk detection error metric. ERDE: early risk detection error.

$$(a)\ ERDE_o(d,k) = \begin{cases} c_{f_p} & if d = positive\ AND\ gt = negative \\ c_{f_n} & if d = negative\ AND\ gt = positive \\ lc_o(k) \cdot c_{tp} & if d = positive\ AND\ gt = positive \\ 0 & if d = negative\ AND\ gt = negative \end{cases}$$

$$(b)\ lc_o(k) = 1 - \frac{1}{1 + exp\ (k - o)}$$

Table 3. Baselines used for comparison with our proposed methods.

Method	ERDE[a]$_5$	ERDE$_{50}$	*F* measure	Precision	Recall
Random	18.51	15.20	0.20	0.12	0.00
All depressed	21.67	15.03	0.23	0.13	1.00
Nondepressed	12.97	12.97	0.00	0.00	0.00
Oracle1	10.38	3.74	1.00	1.00	1.00
Oracle2	11.83	5.30	1.00	1.00	1.00
Oracle3	12.23	6.73	1.00	1.00	1.00
Oracle5	12.59	7.86	1.00	1.00	1.00
Oracle10	12.97	12.97	1.00	1.00	1.00
FHDOB[b]	12.70	10.39	0.55	0.69	0.46
UNSLA[c]	13.66	9.68	0.59	0.48	0.79

[a]ERDE: early risk detection error.

[b]Model B presented by the University of Applied Sciences and Arts Dortmund, Germany (FHDO).

[c]Model A presented by the National University of San Luis, Argentina (UNSL).

We also present the nondepressed method that emits a *nondepression* decision for all subjects. As observed in Table 3, this method scored zero in all effectiveness metrics. The Oracle methods present the results for an oracle that perfectly diagnoses all subjects at the specified chunk (only results for chunks 1, 2, 3, 5, and 10 are displayed). These results prove the difficulty of this task, as the effectiveness metrics (precision, recall, and *F* measure) obtained perfect values, whereas the ERDE metrics showed error values. Oracle10 obtained the same results for nondepression because of the strict penalization of late detection of depression cases (being equivalent to a wrong diagnosis of nondepression).

Finally, the best methods from eRisk 2017 were considered for both ERDE$_5$ and ERDE$_{50}$. The FHDOB method was presented by the Biomedical Computer Science Group from the University of Applied Sciences and Arts Dortmund (Germany). This model employed linguistic metainformation extracted from the subjects' texts and considered classifiers based on bag of words, paragraph vector, LSA, and recurrent neural networks using long short-term memory [38]. The UNSLA method was presented by the Laboratory of Research and Development in Computational Intelligence Research Group from the National University of San Luis (Argentina). This method is based on a semantic representation of documents that explicitly considers the partial information available in different chunks of data, complemented with standard categorization technology. In this case, predictions are based on their own temporal models and other sources of opinion. The LIDIC group considered multiple document representations and different learning algorithms, including RF [39].

An important difference between ERDE$_5$ and ERDE$_{50}$ is that the former promotes methods that emit few yet rapid depression decisions, whereas the latter gives smoother penalties to delays. ERDE$_5$ from FHDOB and ERDE$_{50}$ from UNSLA were used as main baselines for the comparison of our proposed methods.

Evaluation

In this section, we present our main findings for the classification task described, and we discuss the effects of features and the performance for the different proposed models.

The first set of experiments were focused on the singleton model (Table 4).

XSL•FO
RenderX

Table 4. Evaluation results for the singleton model on different feature sets. Writing feature (WF) groups all WFs presented. The values for the best early risk detection error with 0=5 and 0=50 are in italics.

Features	ERDE[a]$_5$	ERDE$_{50}$	*F* measure	Precision	Recall
Cos[b] Text[c]	15.83	13.22	0.31	0.23	0.46
Cos All[d]	16.48	13.62	0.36	0.24	0.67
BM25[e] Text	18.11	16.61	0.26	0.16	0.60
BM25 All	14.36	12.43	0.36	0.32	0.40
LSA[f]	21.60	14.96	0.23	0.13	1.00
Norm[g] LSA	21.34	18.02	0.23	0.13	1.00
LSA stem[h]	23.51	14.70	0.23	0.13	1.00
Norm LSA stem[i]	12.97	12.97	0.00	0.00	0.00
Cos Text + WF	14.09	13.60	0.07	0.33	0.04
Cos All + WF	13.31	12.31	0.20	0.67	0.12
BM25 Text + WF	15.59	14.62	0.29	0.24	0.37
BM25 All + WF	20.49	18.05	0.30	0.18	0.83
Cos BM25 Text + WF	14.15	12.97	0.30	0.38	0.25
Cos BM25 All + WF	13.29	12.97	0.12	0.29	0.08
LSA Cos Text + WF	17.86	12.92	0.29	0.18	0.73
LSA BM25 Text + WF	16.61	12.09	0.27	0.18	0.56
LSA Cos All + WF	19.51	13.46	0.26	0.15	0.90
LSA BM25 All + WF	20.47	14.08	0.24	0.14	0.94
LSA Cos BM25 Text + WF	18.34	12.85	0.28	0.17	0.85
LSA Cos BM25 All + WF	*12.89*	*11.27*	0.34	0.45	0.27
Norm LSA Cos Text + WF	13.35	13.35	0.04	0.20	0.02
Norm LSA BM25 Text + WF	13.58	13.33	0.11	0.17	0.08
Norm LSA Cos All + WF	14.70	14.45	0.11	0.21	0.08
Norm LSA BM25 All + WF	14.55	14.30	0.13	0.22	0.10
Norm LSA Cos BM25 Text + WF	14.60	14.60	0.25	0.25	0.08
Norm LSA Cos BM25 All + WF	13.73	13.48	0.20	0.20	0.08

[a]ERDE: early risk detection error.

[b]Cos: cosine.

[c]Only the text part of the writing is considered.

[d]The whole writing is considered.

[e]BM25: Okapi Best Matching 25.

[f]LSA: latent semantic analysis.

[g]Normalized LSA.

[h]LSA with stemming.

[i]Normalized LSA with stemming.

Initially, we tested the performance of textual and standalone LSA features, finding very low performance on the semantic features, compared with textual features, probably because of the difficulty to capture small textual details relevant to the detection of depression. The normalization of the LSA scores slightly improved the results, and the use of stemming and stop words removal had a negative impact on performance, as shown by the zero values on precision, recall, and *F* measure for normalized LSA with stemming. The best performance among textual features was obtained by BM25 using all textual writing fields (title, info, and text).

Next, we analyzed the performance of textual features combined with the WFs, as defined in Data Analysis. Curiously, BM25 performance worsened as the WFs were included, whereas cosine performance improved. However, the best results were

XSL•FO
RenderX

obtained combining both textual features (cosine and BM25 similarity) with WFs using all textual writing fields.

Subsequently, the 3 feature types were combined (textual, semantic, and writings), and the best results were obtained when the textual similarity metrics (cosine and BM25) used all textual fields, altogether with LSA and WFs. The same set of experiments were executed with normalized LSA and, although the results generally improved, they did not outperform the best value for nonnormalized LSA. Focusing on the best performing singleton model from Table 4, we individually analyzed the results for the different WFs described in Data Analysis to determine these features' behavior. Table 5 shows the results, highlighting the best $ERDE_5$ and $ERDE_{50}$ values. Best singleton model refers to *LSA Cos BM25 All* and the WFs are grouped in the following manner:

- Writing: NWritings, AvgWords, DevWords, MinWords, MaxWords, MedWords.
- TimeGap
- LogTimeGap
- Day
- Week: Weekday, Weekend
- Hour

Regarding a fast early detection (measured through $ERDE_5$), the best performance was obtained just considering the text features, the time gap between writings and the publication days, which closely reflected the conclusions extracted from our data analysis on Section 3 in that a higher tendency to publish during the weekends could be observed in the depressed group. Relatively good results were also obtained combining textual features with the log time gap and writing hours (second best performance). Curiously, the combination of both week group and hour with textual features and time gap led to the worst results of the group. However, the best $ERDE_5$ value from Table 4 was not outperformed by any combination, as each of these features is expected to capture different variables in the writings' behavior.

The best value from the $ERDE_{50}$ metric was obtained by combining text features, both time gap variants and the publication days. Three of these features obtained the best $ERDE_5$ performance in Table 5. In this case, $ERDE_{50}$ outperformed the best value from Table 4, although values are extremely close (11.26 and 11.27, respectively).

The performance values obtained in Tables 4 and 5 do not outperform our baselines, although the results are closer in the case of $ERDE_5$.

Tables 6 and 7 show the performance results in terms of $ERDE_5$ and $ERDE_{50}$ for different dual model configurations. In the case of the dual model, 2 models were trained in parallel: one to detect depression cases (positive) and another to detect nondepression cases (negative). Both tables show a matrix in which the first column indicates the different features considered for the positive model, and the first row provides the features for the negative model (in the same order as the positive features).

Table 5. Evaluation results for classification on different writing features for the best singleton model from Table 4, which combines cosine and Okapi Best Matching 25 textual features for all text fields and latent semantic analysis features. The values for the best early risk detection error with 0=5 and 0=50 are in italics.

WF[a] combinations	$ERDE^b_5$	$ERDE_{50}$	*F* measure	Precision	Recall
BSM[c] + Writing, TimeGap, Hour	17.35	11.39	0.30	0.18	0.85
BSM + Writing, TimeGap, Day	*13.59*	12.12	0.22	0.31	0.17
BSM + Writing, TimeGap, Week	14.77	11.44	0.33	0.25	0.48
BSM + Writing, LogTimeGap, Hour	14.03	13.54	0.12	0.29	0.08
BSM + Writing, LogTimeGap, Day	18.95	12.53	0.27	0.16	0.96
BSM + Writing, LogTimeGap, Week	17.80	12.72	0.28	0.17	0.85
BSM + Writing, TimeGap, Day, Hour	16.14	11.55	0.31	0.21	0.63
BSM + Writing, TimeGap, Week, Hour	19.28	12.85	0.26	0.15	0.94
BSM + Writing, LogTimeGap, Day, Hour	16.86	12.28	0.29	0.18	0.77
BSM + Writing, LogTimeGap, Week, Hour	16.91	12.13	0.29	0.19	0.63
BSM + Writing, TimeGap, LogTimeGap, Day	17.00	*11.26*	0.31	0.19	0.87
BSM + Writing, TimeGap, LogTimeGap, Week	17.85	12.62	0.30	0.18	0.87
BSM + Writing, TimeGap, LogTimeGap, Hour	17.71	12.65	0.28	0.17	0.83
BSM + Writing, TimeGap, LogTimeGap, Hour, Week	16.53	13.47	0.29	0.20	0.52

[a]WF: writing feature.

[b]ERDE: early risk detection error.

[c]BSM: best singleton model.

XSL•FO
RenderX

Table 6. Evaluation results for classification of different feature sets for the dual model (th_w=6). The first column shows features for the positive model, and the first row shows features for the negative model. Positive feature sets are numbered and negative features follow the same numbering. The values for the best early risk detection $error_5$ are in italics. Labels for the algorithms (Roman numerals) are shared for rows and columns.

Features	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
LSA[a] (I)	13.24	12.99	12.99	12.99	12.99	12.99	13.48	13.24	13.48	13.24	29.20	13.24
Norm LSA[b] (II)	13.22	12.97	12.97	12.97	13.47	12.97	13.47	13.22	13.47	13.22	29.43	13.22
LSA stem[c] (III)	13.40	13.15	13.15	13.15	13.15	13.15	13.65	13.40	13.65	13.40	29.36	13.40
Norm LSA stem[d] (IV)	13.22	12.97	12.97	12.97	13.47	12.97	13.47	13.22	13.47	13.22	29.43	13.22
Cos[e] BM25[f] Text[g] + WF[h] (V)	13.22	12.97	12.97	12.97	13.47	12.97	13.47	13.22	13.47	13.22	29.43	13.22
Cos BM25 All[i] + WF (VI)	13.22	12.97	12.97	12.97	13.47	12.97	13.47	13.22	13.47	13.22	29.43	13.22
LSA Cos Text + WF (VII)	13.24	12.99	12.99	12.99	12.99	12.99	13.48	13.24	13.48	13.24	29.20	13.24
LSA BM25 Text + WF (VIII)	13.24	12.99	12.99	12.99	12.99	12.99	13.48	13.24	13.48	13.24	29.20	13.24
LSA Cos All + WF (IX)	12.14	11.89	11.89	11.89	11.89	11.89	12.39	12.14	12.39	12.14	28.35	12.14
LSA BM25 All + WF (X)	13.24	12.99	12.99	12.99	12.99	12.99	13.48	13.24	13.48	13.24	29.20	13.24
LSA Cos BM25 Text + WF (XI)	12.13	*11.88*[j]	*11.88*[j]	*11.88*[j]	*11.88*[j]	*11.88*[j]	12.38	12.13	12.38	12.13	28.34	12.13
LSA Cos BM25 All + WF (XII)	12.73	12.49[j]	12.49[j]	12.49[j]	12.73	12.49[j]	12.98	12.73	12.98	12.73	28.94	12.73

[a]LSA: latent semantic analysis.

[b]Normalized LSA.

[c]LSA with stemming.

[d]Normalized LSA with stemming.

[e]Cos: cosine.

[f]BM25: Okapi Best Matching 25.

[g]Only the text part of the writing is considered.

[h]WF: writing features.

[i]The whole writing is considered.

[j]Statistically significant performance improvements over the best singleton model in Table 4.

Experiments were performed with an extensive number of feature combinations, but we have limited the results displayed on the tables to the most relevant performing features. Focusing on $ERDE_5$ (Table 6), the best results were obtained when using textual features (both cosine and BM25 similarity metrics) only for the text field, LSA and WFs on the positive model, combined with LSA variants or textual features for the negative model. Among the LSA variants, except for plain LSA, normalized LSA, LSA with stemming, and normalized LSA with stemming provide the best performance. The sole use of textual similarity features (feature sets 5 and 6) with any LSA features leads to a best performing model.

We also report on statistical significance using a standard 2-sided *t* test with significance level alpha=.05 when improving performance of the best singleton model on Table 4. Significant improvements (all the *P* values obtained are smaller than 1.21e-14) over the best singleton model were obtained with positive models using both textual features (cosine and BM25) in all fields or just in the text field in combination with semantic and WFs, as well as negative models based on LSA (normalized, stemming, and both). Significant improvement was also achieved using both textual features together with WFs but skipping LSA. This suggests that all the proposed features are required to provide an early risk detection for the identification of depressed subjects, whereas a less complex model achieves better results in identifying nondepressed subjects.

Results for $ERDE_{50}$ (Table 7) are consistent with $ERDE_5$ performance (all the *P* values are smaller than .003), but the optimal value is limited to the positive model with textual features on the text field, LSA and WFs, whereas the negative model only applies LSA with stemming and removing stop words. Other best-performing models from Table 6 obtained the third best performance for $ERDE_{50}$, whereas the second best uses cosine similarity for all text fields, LSA, and writings features for the positive model. It is remarkable that the negative model for both first and second best performance is based only on LSA with stemming and removing stop words. The dual model is able to clearly outperform the best baseline values for $ERDE_5$ and $ERDE_{50}$ from Table 3, with an improvement of 6.5% on $ERDE_5$ and more than 10% improvement on $ERDE_{50}$ over the best-performing state-of-the-art models. Thus, we were able to improve on 2 different and independent best-performing models by employing a single model with two different configurations.

XSL•FO
RenderX

Table 7. Evaluation results for classification of different feature sets for the dual model (th_w=53). The first column shows features for the positive model, and the first row shows features for the negative model. Positive feature sets are numbered and negative features follow the same numbering. The values for the best early risk detection $error_{50}$ are in italics. Labels for the algorithms (Roman numerals) are shared for rows and columns.

Features	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
LSA[a] (I)	10.20	9.95	9.95	9.95	9.95	9.95	10.45	10.20	10.45	9.95	16.18	10.20
Norm LSA[b] (II)	15.46	15.21	12.97	15.21	15.21	15.21	15.71	15.46	15.71	15.46	31.42	5.46
LSA stem[c] (III)	11.19	10.94	10.94	10.94	10.94	10.94	11.44	11.19	11.44	11.19	25.15	11.19
Norm LSA stem[d] (IV)	15.46	15.21	12.97	15.21	15.21	15.21	15.71	15.46	15.71	15.46	31.42	15.46
Cos[e] BM25[f] Text[g] + WF[h] (V)	15.46	15.21	12.97	15.21	15.21	15.21	15.71	15.46	15.71	15.46	31.42	15.46
Cos BM25 All[i] + WF (VI)	15.46	15.21	12.97	15.21	15.21	15.21	15.71	15.46	15.71	15.46	31.42	15.46
LSA Cos Text + WF (VII)	10.20	9.95[j]	9.95[j]	9.95[j]	9.95[j]	9.95[j]	10.45	10.20	10.45	9.95[j]	16.18	10.20[j]
LSA BM25 Text + WF (VIII)	10.20	9.95[j]	9.95[j]	9.95[j]	9.95[j]	9.95[j]	10.45	10.20	10.45	9.95[j]	16.18	10.20[j]
LSA Cos All + WF (IX)	10.41	10.16	9.16	10.16	10.16	10.16	10.66	10.41	10.66	10.16	16.65	10.41
LSA BM25 All + WF (X)	10.32	10.07	10.07	10.07	10.07	10.07	10.57	10.32	10.57	10.07	16.30	10.32
LSA Cos BM25 Text + WF (XI)	10.17	9.93	*8.68*[j]	9.93	9.93	9.93	10.42	10.17	10.42	9.93	17.41	10.17
LSA Cos BM25 All + WF (XII)	13.48	13.23	10.98[j]	13.23	13.23	13.23	13.73	13.48	13.73	13.48	28.94	13.48

[a]LSA: latent semantic analysis.

[b]Normalized LSA.

[c]LSA with stemming.

[d]Normalized LSA with stemming.

[e]Cos: cosine.

[f]BM25: Okapi Best Matching 25.

[g]Only the text part of the writing is considered.

[h]WF: writing features.

[i]The whole writing is considered.

[j]Statistically significant performance improvements over the best singleton model in Table 4.

Discussion

Principal Findings

The main findings of this study are the following: the importance of using WFs in the early detection of MDD, the comparison of the singleton and dual approaches to predict the depression condition, and the improvement of state-of-the-art algorithms, following a time-aware evaluation, obtained by the dual model.

In this paper, we presented 2 methods based on machine learning that exclusively used data from social media networks to provide an early detection of depression cases. The problem was formalized as a classification problem and was addressed using machine learning. We resorted to a features-based approach and designed a collection of features (textual, semantic, and writing) that captured correlations between different aspects of the individuals' writings and depression. The evaluation follows a time-aware approach that rewards early detections and penalizes late detections.

Initially, we present a singleton model based on a single binary classifier and 2 threshold functions (one positive and another negative). However, the results achieved were modest because, to make a final decision, the classifier requires enough evidence to discard one option versus the other, thus causing a delay. The best results for the singleton model were obtained by combining textual and semantic similarity with all the WFs proposed. Note that an individual combination of WFs did not lead to improved results.

Our best-performing method was based on a dual approach, using a machine learning model to detect depressed subjects and another one to identify nondepressed ones. Interestingly, WFs become crucial for the positive model (in charge of detecting depression cases), along with semantic similarity and textual similarity, although limited to the post text field. On the contrary, the negative model (predicting nondepression cases) can follow a much simpler approach based on semantic or textual similarity.

In fact, focusing on $ERDE_{50}$, the optimal value is obtained with the negative model based only on LSA with stemming and removing stop words, without considering any textual similarity or WFs. This may be related with the less strict evaluation of false negatives using this metric.

In comparison with the state-of-the-art detection models, our results showed how the dual model is able to improve performance up to more than 10%. We consider that these results can help in the development of new tools to identify at-risk

XSL•FO
RenderX

individuals, enabling those people suffering from depression to be detected and receive treatment as soon as possible.

Future Work

This study can be extended in several ways. First, we would like to extend the set of features with other document representations. Second, we plan to study different model combinations for our dual approach, with an intense focus on new machine learning algorithms and feature sets. Finally, we plan to evaluate the effectiveness of our models in different environments, such as information technologies or economics.

Acknowledgments

This study was supported by the Ministry of Economy and Competitiveness of Spain and FEDER funds of the European Union (Project TIN2015-70648-P), by the Xunta de Galicia (Singular research center of Galicia, accreditation ED431G/01 2016-2019) and the European Union (European Regional Development Fund).

Conflicts of Interest

None declared.

References

1. Kessler RC, Aguilar-Gaxiola S, Alonso J, Chatterji S, Lee S, Ormel J, et al. The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. Epidemiol Psichiatr Soc 2009;18(1):23-33 [FREE Full text] [Medline: 19378696]
2. Le HN, Boyd RC. Prevention of major depression: early detection and early intervention in the general population. Clin Neuropsychiatry 2006;3(1):6-22 [FREE Full text]
3. World Health Organization. 2013. Comprehensive mental health action plan 2013-2020 URL: https://www.who.int/mental_health/action_plan_2013/en/
4. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Medicine 2006 Nov;3(11):e442 [FREE Full text] [doi: 10.1371/journal.pmed.0030442] [Medline: 17132052]
5. Muñoz RF, Mrazek PJ, Haggerty RJ. Institute of Medicine report on prevention of mental disorders. Summary and commentary. Am Psychol 1996 Nov;51(11):1116-1122. [Medline: 8937259]
6. Halfin A. Depression: the benefits of early and appropriate treatment. Am J Manag Care 2007 Nov;13(4 Suppl):S92-S97 [FREE Full text] [Medline: 18041868]
7. Picardi A, Lega I, Tarsitani L, Caredda M, Matteucci G, Zerella MP, SET-DEP Group. A randomised controlled trial of the effectiveness of a program for early detection and treatment of depression in primary care. J Affect Disord 2016 Dec 01;198:96-101. [doi: 10.1016/j.jad.2016.03.025] [Medline: 27015158]
8. Rost K, Smith JL, Dickinson M. The effect of improving primary care depression management on employee absenteeism and productivity. A randomized trial. Med Care 2004 Dec;42(12):1202-1210 [FREE Full text] [Medline: 15550800]
9. Cameron IM, Cardy A, Crawford JR, du Toit SW, Hay S, Lawton K, et al. Measuring depression severity in general practice: discriminatory performance of the PHQ-9, HADS-D, and BDI-II. Br J Gen Pract 2011 Jul 01;61(588):e419-e426. [doi: 10.3399/bjgp11X583209] [Medline: 21722450]
10. Smarr KL, Keefer AL. Measures of depression and depressive symptoms: Beck Depression Inventory-II (BDI-II), Center for Epidemiologic Studies Depression Scale (CES-D), Geriatric Depression Scale (GDS), Hospital Anxiety and Depression Scale (HADS), and Patient Health Questionnaire-9 (PH-9). Arthritis Care Res 2011 Nov 07;63(S11):S454-S466. [doi: 10.1002/acr.20556]
11. Hamilton M. Rating depressive patients. J Clin Psychiatry 1980 Dec;41(12 Pt 2):21-24. [Medline: 7440521]
12. Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language. use: our words, our selves. Annu Rev Psychol 2003;54:547-577. [doi: 10.1146/annurev.psych.54.101601.145041] [Medline: 12185209]
13. Losada D, Crestani F, Parapar J. eRISK 2017: CLEF Lab on Early Risk Prediction on the Internet: Experimental Foundations. 2017 Presented at: International Conference of the Cross-Language Evaluation Forum for European Languages (eRisk 2017); September 11–14, 2018; Avignon (France) p. 343-360. [doi: 10.1007/978-3-319-65813-1_30]
14. Park M, McDonald D, Cha M. Perception differences between the depressed and non-depressed users in Twitter. 2013 Jul Presented at: International AAAI Conference on Web and Social Media (ICWSM). The AAAI Press; July, 2013; Cambridge, Massachusetts, USA.
15. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. 2013 Jul Presented at: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media; July; 2013; Cambridge, Massachusetts, USA p. 128-137.
16. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J Med Internet Res 2017 Dec 29;19(6):e228 [FREE Full text] [doi: 10.2196/jmir.7215] [Medline: 28663166]
17. Colineau N, Paris C. Talking about your health to strangers: understanding the use of online social networks by patients. New Rev Hypermedia M 2010 Apr;16(1-2):141-160. [doi: 10.1080/13614568.2010.496131]

XSL•FO
RenderX

18. Añel JA. The importance of reviewing the code. Commun ACM 2011;54(5):40-41. [doi: 10.1145/1941487.1941502]
19. Prieto VM, Matos S, Álvarez M, Cacheda F, Oliveira JL. Twitter: a good place to detect health conditions. PLoS One 2014 Jan;9(1):e86191 [FREE Full text] [doi: 10.1371/journal.pone.0086191] [Medline: 24489699]
20. Chunara R, Andrews JR, Brownstein JS. Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. Am J Trop Med Hyg 2012 Jan;86(1):39-45 [FREE Full text] [doi: 10.4269/ajtmh.2012.11-0597] [Medline: 22232449]
21. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
22. Aladağ AE, Muderrisoglu S, Akbas NB, Zahmacioglu O, Bingol HO. Detecting suicidal ideation on forums: proof-of-concept study. J Med Internet Res 2018 Jun 21;20(6):e215 [FREE Full text] [doi: 10.2196/jmir.9840] [Medline: 29929945]
23. Rice SM, Goodall J, Hetrick SE, Parker AG, Gilbertson T, Amminger GP, et al. Online and social networking interventions for the treatment of depression in young people: a systematic review. J Med Internet Res 2014;16(9):e206 [FREE Full text] [doi: 10.2196/jmir.3304] [Medline: 25226790]
24. Balani S, De Choudhury M. Detecting and characterizing mental health related self-disclosure in social media. 2015 Presented at: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems; 2015, April 18-23; Seoul, Republic of Korea p. 1373-1378. [doi: 10.1145/2702613.2732733]
25. Birnbaum ML, Ernala SK, Rizvi AF, De Choudhury M, Kane JM. A collaborative approach to identifying social media markers of schizophrenia by employing machine learning and clinical appraisals. J Med Internet Res 2017 Dec 14;19(8):e289 [FREE Full text] [doi: 10.2196/jmir.7956] [Medline: 28807891]
26. Conway M, O'Connor D. Social media, big data, and mental health: current advances and ethical implications. Curr Opin Psychol 2016 Jun;9:77-82 [FREE Full text] [doi: 10.1016/j.copsyc.2016.01.004] [Medline: 27042689]
27. De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M. Discovering shifts to suicidal ideation from mental health content in social media. 2016 May Presented at: Proceedings of the CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, California, USA p. 2098-2110 URL: http://europepmc.org/abstract/MED/29082385 [doi: 10.1145/2858036.2858207]
28. Guntuku SC, Yaden DB, Kern ML, Ungar LH, Eichstaedt JC. Detecting depression and mental illness on social media: an integrative review. Curr Opin Behav Sci 2017 Dec;18:43-49. [doi: 10.1016/j.cobeha.2017.07.005]
29. Ziemer KS, Korkmaz G. Using text to predict psychological and physical health: a comparison of human raters and computerized text analysis. Comput Hum Behav 2017 Nov;76:122-127. [doi: 10.1016/j.chb.2017.06.038]
30. Nadeem M. arXiv. 2016. Identifying depression on Twitter URL: https://arxiv.org/ftp/arxiv/papers/1607/1607.07384.pdf [accessed 2019-06-04] [WebCite Cache ID 78sMlb9sF]
31. Coppersmith G, Dredze M, Harman C, Hollingshead K, Mitchell M. CLPsych 2015 shared task: Depression and PTSD on Twitter. 2015 Presented at: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; Jun 5, 2015; Dever, Colorado, USA p. 31-39. [doi: 10.3115/v1/W15-1204]
32. Pedersen T. Screening Twitter users for depression and PTSD with lexical decision lists. 2015 Presented at: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, Colorado, USA p. 46-53. [doi: 10.3115/v1/W15-1206]
33. Resnik P, Armstrong W, Claudino L, Nguyen T. The University of Maryland CLPsych 2015 shared task system. 2015 Presented at: Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 5, 2015; Denver, Colorado, USA p. 54-60. [doi: 10.3115/v1/W15-1204]
34. Ophir Y, Asterhan CS, Schwarz BB. Unfolding the notes from the walls: adolescents' depression manifestations on Facebook. Comput Hum Behav 2017 Jul;72:96-107. [doi: 10.1016/j.chb.2017.02.013]
35. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
36. Beck AT, Steer RA, Brown GK. Beck depression inventory II: manual. San Antonio, Texas, USA: Psychological Corporation; 1996.
37. Losada D, Crestani F. A test collection for research on depression and language use. 2016 Presented at: Conference Labs of the Evaluation Forum; September 5-8, 2016; Évora, Portugal p. 28-39. [doi: 10.1007/978-3-319-44564-9_3]
38. Trotzek M, Koitka S, Friedrich C. Linguistic metadata augmented classifiers at the CLEF 2017 task for early detection of depression. 2017 Presented at: Conference Labs of the Evaluation Forum; September 11-14, 2017; Dublin, Ireland.
39. Villegas M, Funez D, Ucelay M, Cagnina L, Errecalde M. LIDIC - UNSL's Participation at eRisk 2017: pilot task on early detection of depression. 2017 Presented at: Conference Labs of the Evaluation Forum; September 11-14, 2017; Dublin, Ireland.
40. Singhal A. IEEE Computer Society Technical Committee on Data Engineering. 2001. Modern information retrieval: a brief overview URL: http://singhal.info/ieee2001.pdf [accessed 2019-06-04] [WebCite Cache ID 78sMzVXBa]
41. Robertson S, Zaragoza H. The probabilistic relevance framework: BM25 and beyond. Foundat Trends Inf Retr 2009;3(4):333-389. [doi: 10.1561/1500000019]
42. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. Discourse Process 1998 Jan;25(2-3):259-284. [doi: 10.1080/01638539809545028]

XSL•FO
RenderX

43. Kaiser HF. The application of electronic computers to factor analysis. Educ Psychol Meas 1960 Apr 01;20(1):141-151. [doi: 10.1177/001316446002000116]
44. Wild F, Stahl C, Stermsek G, Neumann G. Parameters Driving Effectiveness of Automated Essay Scoring with LSA. 2005 Presented at: Proceedings of the 9th International Computer Assisted Assessment Conference (CAA); July, 2005; Loughborough, UK p. 485-494.
45. Willett P, Sparck-Jones K. Readings In Information Retrieval. San Francisco, California, USA: Morgan Kaufmann Publishers, Inc; 1997.
46. Aksoy S, Haralick RM. Feature normalization and likelihood-based similarity measures for image retrieval. Pattern Recognit Lett 2001 Apr;22(5):563-582. [doi: 10.1016/S0167-8655(00)00112-4]
47. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. SIGKDD Explorations. 2009 Nov 16. The WEKA data mining software: an update URL: https://www.kdd.org/exploration_files/p2V11n1.pdf [accessed 2019-06-04] [WebCite Cache ID 78sNK4FHS]
48. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
49. Rifkin R, Aldebaro K. In defense of one-vs-all classification. J Mach Learn Res 2004 Dec;5:101-141 [FREE Full text]

Abbreviations

API: application program interface
BM25: Okapi Best Matching 25
ERDE: early risk detection error
eRisk: early risk prediction on the internet
LSA: latent semantic analysis
MDD: major depressive disorder
RF: random forest
VSM: vector space model
WF: writing feature

Edited by G Eysenbach; submitted 22.10.18; peer-reviewed by M Conway, SK Ernala; comments to author 29.11.18; revised version received 24.01.19; accepted 30.03.19; published 10.06.19.

Please cite as:
Cacheda F, Fernandez D, Novoa FJ, Carneiro V
Early Detection of Depression: Social Network Analysis and Random Forest Techniques
J Med Internet Res 2019;21(6):e12554
URL: http://www.jmir.org/2019/6/e12554/
doi:10.2196/12554
PMID:31199323

©Fidel Cacheda, Diego Fernandez, Francisco J Novoa, Victor Carneiro. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 10.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility of a New Cuffless Device for Ambulatory Blood Pressure Measurement in Patients With Hypertension: Mixed Methods Study

Paula AM Ogink[1*], MSc; Jelske M de Jong[1*], BSc; Mats Koeneman[2], BSc; Mariska Weenk[3], MD; Lucien JLPG Engelen[2]; Harry van Goor[3], MD, PhD; Tom H van de Belt[2], PhD; Sebastian JH Bredie[1], MD, PhD

[1]Department of Internal Medicine, Radboud University Medical Center, Nijmegen, Netherlands

[2]REshape Innovation Center, Radboud University Medical Center, Nijmegen, Netherlands

[3]Department of Surgery, Radboud University Medical Center, Nijmegen, Netherlands

[*]these authors contributed equally

Corresponding Author:
Sebastian JH Bredie, MD, PhD
Department of Internal Medicine
Radboud University Medical Center
Geert Grooteplein 8
Nijmegen,
Netherlands
Phone: 31 243618819
Email: bas.bredie@radboudumc.nl

Related Articles:

Comment in: http://www.jmir.org/2020/4/e15952/

Comment in: http://www.jmir.org/2020/4/e16205/

Abstract

Background: Frequent home blood pressure (BP) measurements result in a better estimation of the true BP. However, traditional cuff-based BP measurements are troublesome for patients.

Objective: This study aimed to evaluate the feasibility of a cuffless device for ambulatory systolic blood pressure (SBP) measurement.

Methods: This was a mixed method feasibility study in patients with hypertension. Performance of ambulatory SBPs with the device was analyzed quantitatively by intrauser reproducibility and comparability to a classic home BP monitor. Correct use by the patients was checked with video, and user-friendliness was assessed using a validated questionnaire, the System Usability Scale (SUS). Patient experiences were assessed using qualitative interviews.

Results: A total of 1020 SBP measurements were performed using the Checkme monitor in 11 patients with hypertension. Duplicate SBPs showed a high intrauser correlation (R=0.86, P<.001). SBPs measured by the Checkme monitor did not correlate well with those of the different home monitors (R=0.47, P=.007). However, the mean SBPs measured by the Checkme and home monitors over the 3-week follow-up were strongly correlated (R=0.75, P=.008). In addition, 36.4% (n=4) of the participants performed the Checkme measurements without any mistakes. The mean SUS score was 86.4 (SD 8.3). The most important facilitator was the ease of using the Checkme monitor. Most important barriers included the absence of diastolic BP and the incidental difficulties in obtaining an SBP result.

Conclusions: Given the good intrauser reproducibility, user-friendliness, and patient experience, all of which facilitate patients to perform frequent measurements, cuffless BP monitoring may change the way patients measure their BP at home in the context of ambulant hypertension management.

(J Med Internet Res 2019;21(6):e11164) doi:10.2196/11164

XSL•FO
RenderX

KEYWORDS

ambulatory blood pressure monitoring; home blood pressure monitoring; cuffless blood pressure device; hypertension

Introduction

An elevated blood pressure (BP) is a major risk factor for cardiovascular morbidity and mortality [1]. BP is, however, a highly variable vital parameter, and circumstances under which measuring takes place may influence the result extensively [2-4]. Compared to office BP measurement, home BP measurement predicts cardiovascular risk better [5-8]. The predictive value increases progressively with the number of home measurements [9]. Thus, to improve assessment of BP for diagnosis and management of high BP, the BP needs to be monitored frequently, preferably at home.

For patients with hypertension, home blood pressure monitoring (HBPM) is easy to perform, reliable, and reproducible [10]. Therefore, it is recommended as a routine component of BP monitoring in the American Society of Hypertension and the American Heart Association guidelines [5]. The use of HBPM improves hypertension control and the associated outcomes [11-14]. At home, BP is predominantly measured with an oscillometric BP monitor, which uses an arm cuff. In daily practice, patients use different types of BP monitors, which, in most cases, have been validated according to the international standard and have been checked by their provider [15]. However, in a cross-sectional study by Ruzicka et al, 30% of home BP monitors were found to be inaccurate with use of a stringent criterion of 5 mm Hg difference between measurements. Using a different threshold for accuracy (difference of more than 10 mm Hg), 16 of 210 (8%) HBP monitors were inaccurate for systolic BP (SBP) and 18 (9%) were inaccurate for diastolic BP [16]. Although an automatic BP monitor is relatively easy to use and inexpensive [5], measuring BP is time-consuming and may be perceived as inconvenient. In addition, various factors may influence the accuracy of measuring BP, such as discomfort by inflation of the cuff [17] and inappropriate cuff size and cuff position at the arm in relation to the heart level [18].

A new technique has been developed to measure SBP fast and easy without the use of a cuff, which is applied in new devices such as the Checkme Pro Health Monitor (Shenzhen Viatom, China). An algorithm calculates the SBP based on the pulse transit time determined by the peripheral capillary oxygen saturation (SpO_2) measurement (an estimate of the amount of oxygen in the blood and is the percentage of oxygenated hemoglobin compared to the total amount of hemoglobin in the blood), the electrical electrocardiogram (ECG) signal, and the individual's arterial compliance [19]. For the latter, the cuffless device needs a calibration procedure, which is developed by entering a classically obtained SBP. Calibration needs to be repeated monthly. An SBP measurement with the Checkme monitor takes less than 30 seconds, which could increase the willingness of patients to measure their SBP more frequently. Cuffless BP measurement is an emerging technique, which may lead to an increased patient compliance in measuring BP at home and promoting a larger number of BP results for hypertension management. Moreover, the Checkme monitor captures a single-lead ECG and photoplethysmogram signal in the same 30 seconds.

Recently, we evaluated the validity of the Checkme monitor's SBP results by using criteria of the European Society of Hypertension for validating new BP devices [20]. This validation protocol may be considered inadequate, as it lacks consensus about the quality of the BP measurement to be used for calibrating cuffless devices. However, results obtained with the Checkme monitor were promising over a wide range of BP levels [20]. A recent study showed promising results of vital parameter measurements in an inpatient setting [21]. Since the previous study by Schoot et al was performed under demanded controlled circumstances, there was a need to evaluate the performance of the Checkme monitor in an uncontrolled home setting.

To assess feasibility of the Checkme monitor in an outpatient setting, we studied the performance, user-friendliness, and patient experience of the Checkme me in participants' home settings.

Methods

Research Design, Setting, and Participants

We conducted a pilot study using a mixed method approach. To determine performance of the SBP measurement, we systematically assessed the reproducibility of the Checkme monitor, its comparability to a home BP monitor, and the performance of daily vital measurements with the Checkme monitor using video analysis. A System Usability Scale (SUS) questionnaire was used to determine the user-friendliness. Patient experience was assessed using a semistructured interview following the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) framework. Participants were recruited from the hypertension outpatient clinic of an academic hospital in The Netherlands from April 2017 to May 2017. The institutional review board approved the study (ID: 2017-3241). All participants provided signed informed consent after written and verbal information were obtained.

Inclusion and Exclusion Criteria

Patients were considered eligible if they were receiving medical treatment for high BP, accustomed to home BP measurements with their own blood pressure monitor, of age ≥ 18 years, and had the cognitive ability to understand instruction and perform measurements correctly after instruction. Patients with a pacemaker and pregnant women were excluded.

The Checkme Pro Health Monitor

We evaluated the Checkme Pro Health Monitor, which measures SBP without the use of a cuff. The device also measures a one-lead ECG, heart rate, and SpO_2 in one measurement called "daily check." The method of measurement of these vitals by the Checkme monitor is shown in Figure 1. The right thumb, right middle finger, and left palm are placed on the ECG sensors.

XSL•FO
RenderX

The right index finger is placed on the built-in SpO_2 sensor. To increase accuracy of the results, the device needs to be held steady at the heart level during the measurement. The latest version of the Checkme monitor, used in this study, is cleared by the US Food and Drug Administration (FDA) for measuring these vitals (FDA 510k release: K150869; Device Name: CheckMe Pro Health Monitor; Regulation Number: 21 CFR 870.2300; Regulation Name: Cardiac Monitor Including Cardiotachometer and Rate Alarm; Regulatory Class: Class II; Product Code: MWI on November 6, 2015) and complies with the Conformité Européene (CE) marking medical devices directive (CE certificate was issued on behalf of TüV Rheinland LGA Products GmbH notified body [CE 0197] on Viatom's Health Monitor models "Checkme Pro, Plus, Pod, and Lite" [standard MDD 93/42/EEC, Annex II; Certificate HD60107767 0001; April 27, 2016]).

Figure 1. Demonstration of a systolic blood pressure measurement using the Checkme Pro Health Monitor. The right thumb, middle finger, and left palm are placed on the electrocardiogram sensors. The right index finger is placed on the built-in SpO_2 (peripheral capillary oxygen saturation) sensor. Systolic blood pressure measurement is performed in less than 30 seconds, holding the device steady at heart level.

Study Procedures

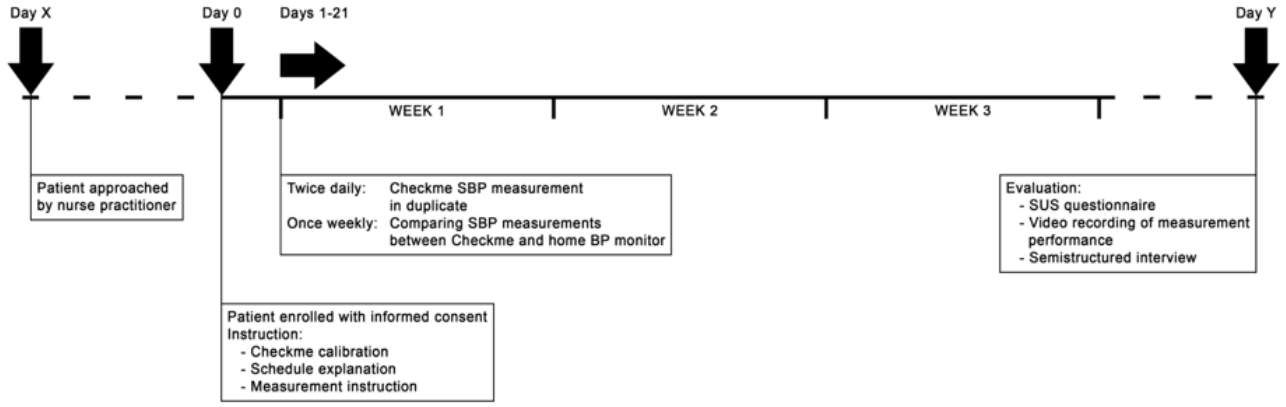
The study timeline and procedures are shown in Figure 2. Two trained researchers instructed the participants on the study procedures and how to perform the SBP measurement with the Checkme monitor, and checked the way they performed their regular home SBP measurement using their own home BP monitor. Before the start of the study, the Checkme monitor was calibrated for SBP measurement according to manufacturer's instructions. To determine a reference SBP to calibrate the Checkme monitor, a standard duplicate BP measurement was performed using a validated automatic BP monitor at the outpatient clinic (Vital Signs Monitor 300 series, Welch Allyn, Skaneateles Falls, NY) after an initial 5 minutes of rest. The participants performed the measurements with the Checkme monitor in duplicate twice daily, in the morning and evening, for a period of 3 weeks. They were instructed to perform the second measurement immediately after the first measurement under the same circumstances.

In addition, the participants performed regular BP measurement once weekly with their own home BP device. Participants were asked to perform one duplicate SBP measurement with the Checkme monitor and one duplicate home BP measurement using their own conventional BP monitor, in a random order.

After 3 weeks, the correct use of the Checkme monitor by the patient was checked with a video recording of the SBP measurement. The user-friendliness was assessed using the SUS questionnaire, and the patient's experience was determined with a semistructured interview.

XSL•FO
RenderX

Figure 2. Timeline of the study procedures. Dotted lines represent a variable time of 0-5 days between day X and day 0, and 0-5 days between the end of the study period and day Y. BP: blood pressure; SBP: systolic blood pressure; SUS: System Usability Scale.



Analysis of Reproducibility

To obtain the intrauser reproducibility of a duplicate SBP measurement, two values of one duplicate measurement were correlated and the level of variation was categorized as <5 mm Hg, <10 mm Hg, and >15 mm Hg. Both the Checkme and home BP monitors were tested for reproducibility. The paired SBP measurements with the Checkme monitor and home BP monitor were correlated to obtain comparability. The mean difference and level of variation in SBP measurements between the Checkme and home BP monitors were calculated. In addition to the paired measurements, the means of all SBP measurements with the Checkme and home BP monitors were correlated. For each participant, all SBP values measured with the Checkme monitor, their home BP monitor, and the hospital monitor were plotted in a diagram to show the variation of SBP over time measured with different devices.

Analysis of the Correct Use of Checkme by the Patient

Two researchers independently assessed the use of Checkme by the patient, by checking the video-recorded measurements with a scoring sheet based on the principles of Gelbart et al [22] and Van Der Heide et al [23]. Thirteen steps were distinguished for the SBP measurement with the Checkme monitor. All items for the use of the Checkme monitor were categorized as “badly performed” or “not done” (0 points), “suboptimal” or “too late” (1 point), and “perfectly done” (2 points). Findings were compared and discussed until a consensus was reached.

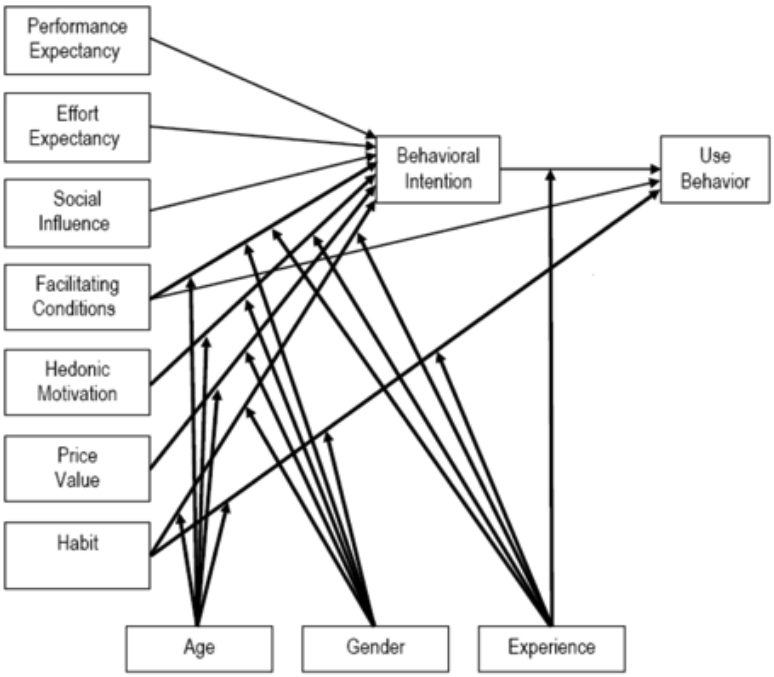
Analysis of Patient Experience

The semistructured interviews following the UTAUT2 framework [24] (Figure 3) were conducted in Dutch with the participants. The UTAUT2 framework consists of four major themes: performance expectancy, effort expectancy, social influence, and facilitating conditions. Performance expectancy is defined as the degree to which an individual believes that using the Checkme monitor will help him/her. Effort expectancy is defined as the degree of ease associated with the use of the Checkme monitor. Social influence is defined as the degree to which an individual perceives that significant others believe he/she should use Checkme. Facilitating conditions are defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the Checkme monitor [25].

Interviews were audio recorded and transcribed verbatim using qualitative data analysis software (ATLAS.ti 7.1, Scientific Software Development GmbH, Berlin, Germany). Transcripts were independently analyzed by two investigators to identify barriers, facilitators, and positive and negative effects of the use of the Checkme monitor. Findings were discussed until a consensus was achieved. The barriers and facilitators were rewritten into general statements and subdivided according to the themes of the UTAUT2 interview framework. The magnitude of each statement was determined by the number of interviews the statement was mentioned in. A validated Dutch translation [26] of the SUS questionnaire [27] on the usability of the Checkme monitor was used to determine the user-friendliness, scored between 0 and 100, as described by Brooke et al [27]. The interpretation of the SUS score was in accordance with that provided by Bangor et al [28]. A score above 90.9 was considered “best imaginable,” a score above 85.5 was considered “excellent,” a score above 71.4 was considered “good,” a score above 50.9 was considered “sufficient,” and a score below or equal to 50.9 was considered “poor.”

XSL•FO
RenderX

Figure 3. Interview framework of the Unified Theory of Acceptance and Use of Technology 2.



Statistical Analyses

All statistical analyses were performed using IBM SPSS, version 22.0 (IBM Corp, Armonk, NY). Normally distributed data were presented as mean and SD. Descriptive statistics were presented as median and quartiles in case of nonnormally distributed data. Differences were tested using a *t* test in case of normal distribution of the data and the nonparametric Wilcoxon test in case of nonnormally distributed data. Correlations were calculated with the Spearman rank correlation coefficient.

Results

User Statistics

One of 12 participants enrolled in the study was excluded from participation and analysis due to repeated failure of BP calibration. One participant did not own a BP monitor and visited the hospital for weekly BP measurements. Average instruction time was 20-40 minutes. The characteristics of the 11 participants who completed the study period are summarized in Table 1. The SBP readings of the participants' home BP monitors strongly correlated with those of the automatic hospital BP monitor at baseline (R=0.88, P<.001). Eleven participants performed a total of 1020 measurements with the Checkme monitor. In 209 measurements (20.4%), the Checkme monitor was not able to measure the SBP. The success rate for SBP measurement varied among participants, with a mean success of 71%, ranging from 42% to 100%.

XSL•FO
RenderX

Table 1. Participant characteristics (N=11).

Characteristics	Study population
Gender, n (%)	
Female	4 (36)
Male	7 (64)
Ethnicity, n (%)	
Caucasian	10 (91)
Black	1 (9)
Age (years), mean (SD)	57 (11.5)
Systolic BP[a] (mm Hg), mean (SD)[b]	140.7 (13.7)
Diastolic BP (mm Hg), mean (SD)[b, c]	86.3 (11.0)
Use of BP-lowering medication, n (%)	9 (82)
Use of home monitor, n (%)	10 (91)
Brand, n[d]	
Withings	3
Microlife	2
Omron	1
Beurer	1
A&D Medical	1
Medion	1
Cresta	1
Frequency per month, mean (SD)	6.8 (6.2)

[a]BP: blood pressure.

[b]BP measured by trained investigator with a Welch Allyn Automatic BP monitor at day 0.

[c]Data shown for only 10 patients, because of the lack of diastolic BP data in one patient.

[d]N=10.

Reproducibility

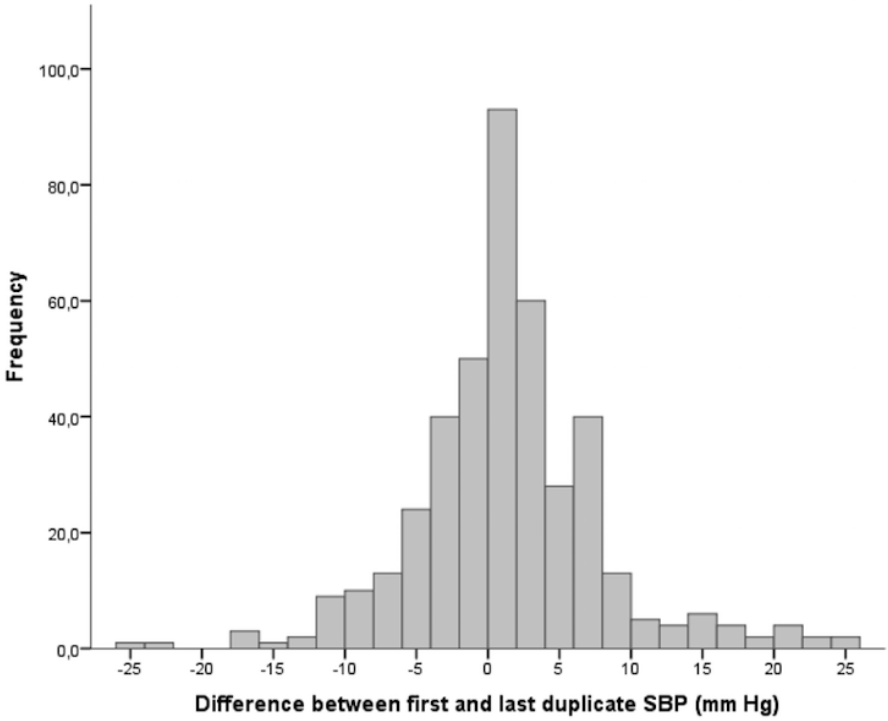
The paired results of duplicate SBP measurements of the Checkme monitor correlated well over the whole range of BP levels (R=0.86, P=.001). Of the 420 complete duplicate SBPs, paired results of 374 (89%) duplicates varied within 10 mm Hg, of which 286 (68% of total) varied within 5 mm Hg. Paired results of 22 (5%) duplicate SBPs varied more than 15 mm Hg. Variations of the paired results of duplicates are shown in Figure 4. Of the 22 duplicates with a difference of more than 15 mm Hg, 11 were obtained from only two participants. The paired results of duplicate SBP measurements with the home BP monitors correlated strongly (R=0.91, P<.001). Of the 40 complete duplicate SBPs, 38 (95%) varied within 10 mm Hg, of which 27 (67% of total) varied within 5 mm Hg. No measurement exceeded a variation of 15 mm Hg. For each participant, all SBP values measured with the Checkme monitor (twice daily), the home BP monitor (once weekly), and the hospital monitor (once) are plotted in Figure 5.

The measurements of the Checkme and home BP monitors correlated weakly (R=0.47, P=.007). The mean results of the paired measurements with the Checkme monitor were 0.55 mm Hg (SD 12.32) higher than those of the home BP monitors. Of the 32 paired SBP values, there was a difference between the Checkme and home BP monitors of <5 mm Hg in 7 pairs (22%), <10 mm Hg in 18 pairs (56%), and <15 mm Hg in 26 pairs (81%).

The mean SBP of both devices over 3 weeks correlated strongly (R=0.75, P=.008). The Checkme monitor had a mean systematic difference of 0.26 mm Hg (SD 7.66) for mean SBP over 3 weeks compared to the home BP monitors. In addition, 36.3% of the mean SBP measurements with the Checkme and home BP monitors varied within 5 mm Hg and 90.9% varied within 10 mm Hg. No measurement exceeded a variation of 15 mm Hg.

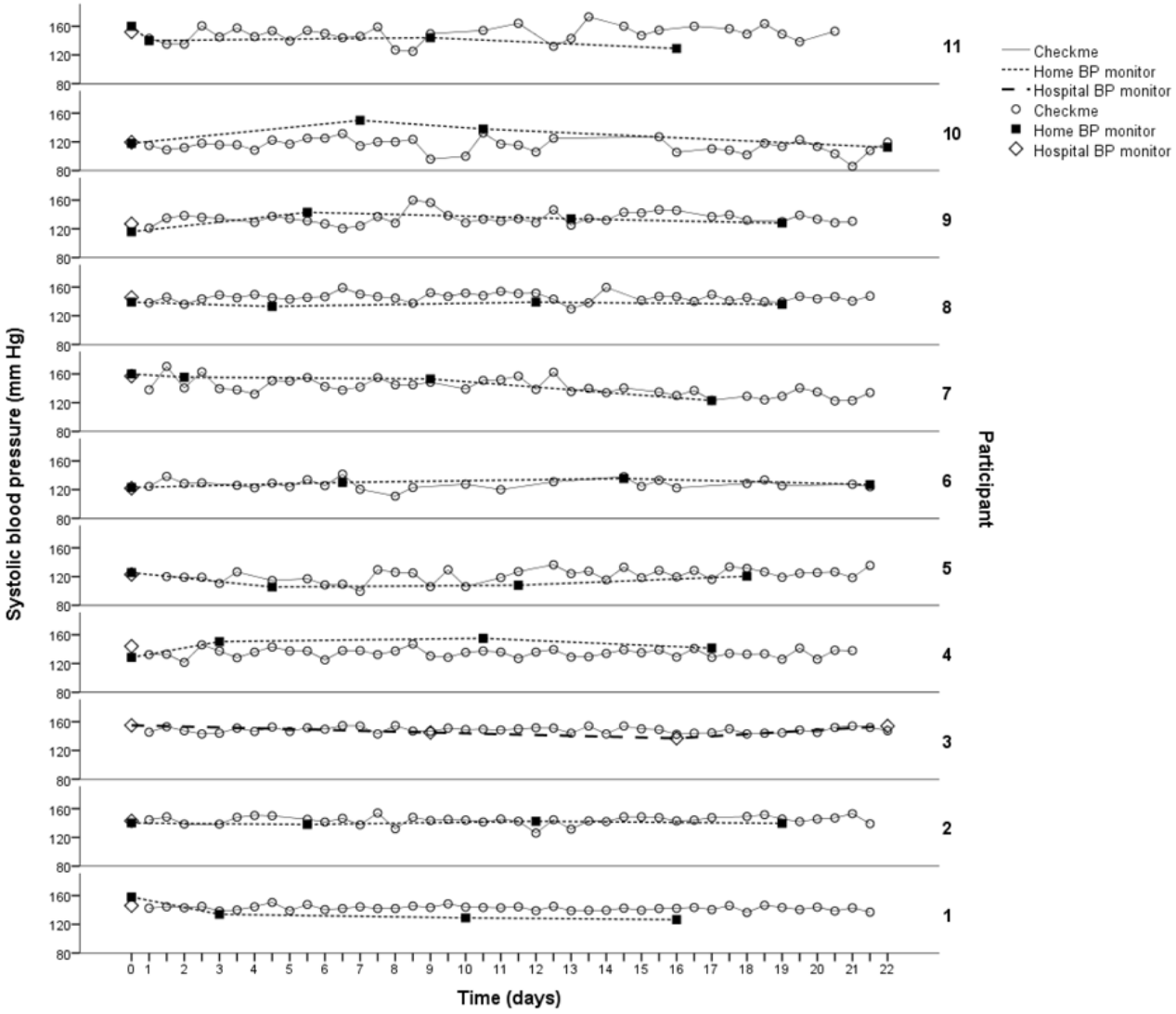
XSL•FO
RenderX

Figure 4. Frequency of the difference within duplicate SBP measurements. SBP: systolic blood pressure.



XSL•FO
RenderX

Figure 5. Systolic blood pressure during follow-up for each participant, measured with Checkme and home BP monitor or hospital BP monitor. BP: blood pressure.



Correct Use of the Checkme Monitor

In total, 36.4% (n=4) of the participants performed the measurement with the Checkme monitor correctly ("well done") in all 13 items, 54.5% (n=6) performed the measurement with one mistake, and 9.1% (n=1) performed the measurement with two mistakes. No participants made more than two mistakes. The most frequent mistake was not keeping the Checkme monitor at heart level (5/11). During video recording, 1 of the 11 participants did not receive a valid SBP result. The hands of this participant were shaking due to the side effects of the medication, and he did not find support by resting his arms on the table. However, at home, this participant achieved a valid SBP in 80 of 104 measurements (77%).

Patient Experience

Interviews lasted for 15-35 minutes. All perceived barriers and facilitators could be subdivided into one of the five themes of the UTAUT2 interview framework. Most significant barriers and facilitators are described here, and all barriers and facilitators are summarized in Table 2.

Performance expectancy could be divided into two subsequent topics: the features and possibilities of the Checkme monitor and the measurement results produced by the Checkme monitor. For the possibilities of the Checkme monitor, six participants perceived the inability of Checkme to measure DBP as a barrier. They considered DBP to be as important as SBP. At times, the Checkme monitor did not report a result for SBP, which was a barrier for seven participants. This was sometimes perceived as bothersome, since the cause was unclear and the measurement had to be repeated. Another barrier mentioned by three participants was the occasional big difference in SBP measured with the Checkme monitor compared to the home or hospital BP monitor. This reduced their trust in Checkme's performance. One participant, on the other hand, reported that the Checkme and home BP monitors correlated well, which increased her trust in the Checkme monitor.

XSL•FO
RenderX

Table 2. Barriers and facilitators for use of the Checkme monitor and the number of interviews these were mentioned in, according to the themes of the Unified Theory of Acceptance and Use of Technology 2 interview framework and subsequent topics.

Variable	Barrier	Facilitator
Performance expectancy	16	1
Possibilities of Checkme		
The device measures only SBP[a]	6	0
Outcomes of Checkme		
At times, the Checkme did not report a result for SBP and/or SpO_2[b]	7	0
The big/small difference between the home BP[c] monitor and the Checkme leads to less/more trust in the Checkme	3	1
Effort expectancy	7	22
Performing measurements		
The Checkme is easy to use	0	11
With the Checkme, a measurement is quickly performed	1	4
BP can be measured with the Checkme without the use of an arm cuff	0	3
Daily check cannot be performed with cold hands	1	0
Design of Checkme		
The Checkme is small and can be taken everywhere	0	4
The Checkme does not have a backlight in the touch screen	2	0
The Checkme is not a standard BP monitor, which decreases trust in results	2	0
The font size of the results screen is very small	1	0
Social influence	0	1
Measuring BP with the Checkme can be done without any help	0	1

[a]SBP: systolic blood pressure.

[b]SpO_2: peripheral capillary oxygen saturation.

[c]BP: blood pressure.

Effort expectancy was defined through two topics: performing measurements and the design of the Checkme monitor. For performing measurements, all participants considered the Checkme monitor easy to use and four could quickly perform a measurement. Three participants perceived the Checkme monitor to be a facilitator that can measure SBP without the use of an arm cuff, mostly because the arm cuff on their home BP monitor was uncomfortable. Regarding the design of the Checkme monitor, the most significant facilitator perceived was the small size of the device (n=4).

In addition, three participants thought the Checkme monitor would be unsuitable for the elderly, because of their decreased fine motor skills. Automatic synchronization of results was preferred by eight participants, either to their medical record or an online app, to be able to monitor the results of medication, diet, and physical activities and discuss the results with their doctor. Two other suggestions were increasing the font size and addition of a backlight to the screen.

Five participants wanted to use the Checkme monitor in the future instead of their own home BP monitor. Three other participants wanted to use the Checkme monitor in the future only on certain conditions, for example, if the device reports reliable results. The remaining three participants did not want to use Checkme in the future. Eight participants would recommend the Checkme monitor to other patients. Participants gave the monitor a median score of 7.5 (interquartile range: 5.0-8.0) on a scale of 1 to 10, with individual scores ranging from 1.0 to 9.0. The mean SUS score was 86.4 (SD 8.3) with a range of 72.5-97.5, which indicates high user-friendliness.

Discussion

Principal Findings

This study provides new insights about the use, performance, and patient experience of an FDA-approved cuffless BP-measuring device in patients who are used to measuring their blood pressure at home. Adequate intrauser reproducibility of cuffless SBP measurement was observed in the majority of participants, and the Checkme monitor was well adopted in the home setting. Patients indicated an increased willingness to take their BP measurement because of its ease of use. Thus, the large variety of cuff-based BP monitors currently used by patients in home BP monitoring does not necessarily serve as a gold standard to compare new devices for home monitoring. In addition, the easily obtained large number of the Checkme SBP measurements may provide a better picture of the actual BP variation over time, which can be easily missed by taking only one home measurement every week.

XSL•FO
RenderX

The SBP measured by the Checkme monitor was comparable to that measured by an in-hospital reference monitor, with a mean difference of 2.6 (SD 12.1) mm Hg [20]. Other studies compared different cuffless SBP-measuring devices to a reference monitor. Poon et al [29] described a mean difference of 0.6 (SD 9.8) mm Hg, and Boubouchairopoulou et al [30] found a mean difference of 3.2 (SD 6.7) mm Hg, which is, to a great extent, comparable to the results of this study. In addition, this study showed that currently, an unrestricted range of BP monitors from different manufacturers are being used for home BP monitoring. Both the variety of home BP monitors and the uncontrolled use in a home setting may contribute to the observed differences. Although home BP devices should ideally be on the list of validated monitors and the circumstances under which measurements are taken should be standardized, the added value of home BP monitoring is the increasing number of results, not the absolute value of each of them [31]. Robust hypertension management is based on the average of a large series of BP measurements rather than a single clinic measurement [31]. The majority of participants could produce a series of valid measurement results, and most of the unsuccessful SBP measurement attempts were observed in a small number of participants. Failure to produce valid SBP readings may be caused by several factors. Since the cuffless technique requires an ECG and SpO_2 signal to produce an SBP result, factors influencing ECG and SpO_2 accuracy may lead to unsuccessful measurements with no SBP results. These factors include poor perfusion (cold fingers) and skin color [32,33]. Performance-related factors such as moving during the measurement or applying too much pressure on the sensors [34] may disturb the SpO_2 signal and thereby influence the SBP result, which suggests that proper user instructions are necessary. Technical factors such as system failure, incorrect calibration procedure, or imperfections in the algorithm may also influence the BP results.

Another issue of the cuffless BP measurement technique is the need of a classic reference BP measurement to calibrate the calculating algorithm for individual vascular compliance. An international standard for this calibration procedure is still lacking in existing protocols for new BP device validation [35]. Schoot et al [20] recently performed a pragmatic validation study with the Checkme monitor by using standardized measurement conditions, which revealed promising results.

Despite its easy-to-use concept, accurate self-measuring with the Checkme monitor was not completely adequate after a single instruction at the start of this study. This phenomenon was also observed in studies with conventional BP monitors, which reported that 52%-65% of patients missed at least one step of the BP measurement process [36,37]. Milot et al found that only 18% of patients performed the classic BP measurement with cuff with excellence [38], and Wagner et al found that none of the participants performed BP measurement correctly [39]. Compared with these observations in classical cuff-based BP measurement, the correct use of the Checkme monitor in our study was much better. The only observed mistake was not holding the device at the heart level, which has a minor effect on the SBP result [20]. Other mistakes concerning the use of the Checkme monitor were not observed. This is in contrast to the observation during classic BP measurement, in which various other errors, with respect to cuff usage and position, can occur. An important finding of the present study is that user instruction needs attention, both at the start and during long-term use, to increase the quality of BP readings. This may be achieved by optimizing the patient instruction by using the protocol described by Mengden et al as a guide [40] or using video instructions [41].

Performance expectancy and effort expectancy were most mentioned in the interviews, and only a few barriers and facilitators were mentioned on social influence and facilitating conditions. This can be explained by the short follow-up period and the fact that Checkme is a new unknown device. The two most prominent issues of performance expectancy for participants are that the Checkme monitor only measures SBP and not DBP, and the monitor sometimes fails to produce a valid SBP result. The cuffless BP measurement technique is currently unable to determine DBP accurately. However, it is internationally accepted that SBP is the primary target in managing cardiovascular risk in most patient groups, except in elderly people [42]. Some participants also suggested improving the design of the Checkme monitor. Although requirements for medical devices are dictated by appropriate legislative bodies such as the European directive (93/42/EEC, the Medical Devices Directive) for the European devices, five requirements for home monitoring with wearable sensors have been described by Korhonen et al: reliability and durability, looks and unobtrusiveness, user identification, communication, and zero maintenance and fault recovery [43]. Cuff-based home BP monitors meet the first and last requirement. The Checkme monitor formally meets the second, third, and fourth criteria with its size, personal user profiles, and ability to share readings, respectively. This study also provides new information about the reliability through its evaluation of intrauser reproducibility of the Checkme monitor in home monitoring. The memory capacity of the Checkme monitor and the ability to automatically share saved readings bypasses the imprecision of self-reported BP readings by patients, which appeared to range from 0% to 100% in a study on 30 patients with hypertension [44]. Further, it enables physicians to intervene and adjust medication, since patients are often not able to interpret the readings of SBP correctly [45,46].

Strengths and Limitations

The strength of this pilot study is that SBP was measured by the cuffless Checkme device in a home-based setting. This is the first study in which the feasibility, usability, and acceptability of a cuffless BP measuring device in a home monitoring setting were assessed using a mixed method study design. We obtained a large number of home measurements in the morning and evening, as recommended by the European Society of Hypertension/European Society of Cardiology [47] for well-instructed patients who were involved in self-management. Patient experience and performance were evaluated by a widely used and reliable questionnaire to determine the feasibility of different products, and all interviews followed an interview guide derived from a well-known interview framework [28,48]. A weakness of this study was the small study sample, which may not guarantee complete

XSL•FO
RenderX

saturation in all qualitative aspects. The relative short follow-up may also have led to an incomplete user experience. In addition, the fact that medication use may instantly influence measurements could induce device-independent differences. However, the blood pressure results that were compared were based on daily averages or were time-related (measurement with both devices at the same time). To confirm the current results, this study should be repeated on a larger scale. Some adjustments of the methods need to be taken into consideration, including more explicit user instructions and a restricted set of validated home BP monitors as a reference. Future studies should focus on the cause and mechanisms of failures to measure SBP with Checkme in some patients or settings. Furthermore, an international validation protocol for the calibration procedure of a cuffless device is needed. The Institute of Electrical and Electronics Engineers Standard for Wearable Cuffless Blood Pressure Measuring Devices [49] should incorporate a norm for such devices.

Implications for Practice

It is highly possible that the use of cuffless SBP devices will become part of common practice in hypertension and cardiovascular risk management in the near future. Therefore, health care professionals should be aware of this development and familiarize themselves with the specific characteristics of these devices. They could explore possibilities such as smart data analysis and connectivity with electronic health records. Patients and their relatives should not hesitate to discuss the possibilities in home monitoring with their health care professionals. If they start using these devices, it may provide them with better insight into their health status and recovery, with minimum effort.

Conclusions

As confidence in BP measurement results continues to increase, and if international consensus on the calibration process is reached, cuffless BP monitoring devices such as the Checkme monitor may change the way patients measure their BP at home, in the context of ambulant hypertension and cardiovascular risk management. A major advantage of the Checkme monitor in addition to the current use of cuff-based BP home monitors is that the former stimulates the patient in taking a larger number of BP readings because of its easy-to-use design.

Conflicts of Interest

None declared.

References

1. Wilson P. Established risk factors and coronary artery disease: the Framingham Study. Am J Hypertens 1994 Jul;7(7 Pt 2):7S-12S. [doi: 10.1093/ajh/7.7.7s] [Medline: 7946184]
2. Verberk W, Kroon AA, Kessels AGH, de Leeuw PW. Home blood pressure measurement: a systematic review. J Am Coll Cardiol 2005 Sep 06;46(5):743-751 [FREE Full text] [doi: 10.1016/j.jacc.2005.05.058] [Medline: 16139119]
3. Kleinert H, Harshfield GA, Pickering TG, Devereux RB, Sullivan PA, Marion RM, et al. What is the value of home blood pressure measurement in patients with mild hypertension? Hypertension 1984;6(4):574-578. [Medline: 6235190]
4. Parati G, Pickering TG. Home blood-pressure monitoring: US and European consensus. Lancet 2009 Mar 14;373(9667):876-878. [doi: 10.1016/S0140-6736(09)60526-2] [Medline: 19286071]
5. Pickering T, Miller NH, Ogedegbe G, Krakoff LR, Artinian NT, Goff D, American Heart Association, American Society of Hypertension, Preventive Cardiovascular Nurses Association. Call to action on use and reimbursement for home blood pressure monitoring: a joint scientific statement from the American Heart Association, American Society of Hypertension, and Preventive Cardiovascular Nurses Association. J Cardiovasc Nurs 2008;23(4):299-323. [doi: 10.1097/01.JCN.0000317429.98844.04] [Medline: 18596492]
6. Asayama K, Ohkubo T, Kikuya M, Metoki H, Hoshi H, Hashimoto J, et al. Prediction of stroke by self-measurement of blood pressure at home versus casual screening blood pressure measurement in relation to the Joint National Committee 7 classification: the Ohasama study. Stroke 2004 Oct;35(10):2356-2361. [doi: 10.1161/01.STR.0000141679.42349.9f] [Medline: 15331792]
7. Niiranen T, Hänninen MR, Johansson J, Reunanen A, Jula AM. Home-measured blood pressure is a stronger predictor of cardiovascular risk than office blood pressure: the Finn-Home study. Hypertension 2010 Jun;55(6):1346-1351. [doi: 10.1161/HYPERTENSIONAHA.109.149336] [Medline: 20385970]
8. Breaux-Shropshire T, Judd E, Vucovich LA, Shropshire TS, Singh S. Does home blood pressure monitoring improve patient outcomes? A systematic review comparing home and ambulatory blood pressure monitoring on blood pressure control and patient outcomes. Integr Blood Press Control 2015;8:43-49 [FREE Full text] [doi: 10.2147/IBPC.S49205] [Medline: 26170715]
9. Niiranen T, Johansson JK, Reunanen A, Jula AM. Optimal schedule for home blood pressure measurement based on prognostic data: the Finn-Home Study. Hypertension 2011 Jun;57(6):1081-1086. [doi: 10.1161/HYPERTENSIONAHA.110.162123] [Medline: 21482956]
10. Mejia A, Julius S. Practical utility of blood pressure readings obtained by self-determination. J Hypertens Suppl 1989 May;7(3):S53-S57. [Medline: 2668465]

XSL•FO
RenderX

11. Omboni S, Gazzola T, Carabelli G, Parati G. Clinical usefulness and cost effectiveness of home blood pressure telemonitoring: meta-analysis of randomized controlled studies. J Hypertens 2013 Mar;31(3):455-67; discussion 467. [doi: 10.1097/HJH.0b013e32835ca8dd] [Medline: 23299557]
12. Duan Y, Xie Z, Dong F, Wu Z, Lin Z, Sun N, et al. Effectiveness of home blood pressure telemonitoring: a systematic review and meta-analysis of randomised controlled studies. J Hum Hypertens 2017 Dec;31(7):427-437. [doi: 10.1038/jhh.2016.99] [Medline: 28332506]
13. Uhlig K, Patel K, Ip S, Kitsios GD, Balk EM. Self-measured blood pressure monitoring in the management of hypertension: a systematic review and meta-analysis. Ann Intern Med 2013 Aug 06;159(3):185-194. [doi: 10.7326/0003-4819-159-3-201308060-00008] [Medline: 23922064]
14. Pickering T, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, et al. Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Circulation 2005 Feb 08;111(5):697-716. [doi: 10.1161/01.CIR.0000154900.76284.F6] [Medline: 15699287]
15. Tislér A, Dunai A, Keszei A, Fekete B, Othmane TEH, Torzsa P, et al. Primary-care physicians' views about the use of home/self blood pressure monitoring: nationwide survey in Hungary. J Hypertens 2006 Sep;24(9):1729-1735. [doi: 10.1097/01.hjh.0000242396.15097.f3] [Medline: 16915021]
16. Ruzicka M, Akbari A, Bruketa E, Kayibanda JF, Baril C, Hiremath S. How Accurate Are Home Blood Pressure Devices in Use? A Cross-Sectional Study. PLoS One 2016;11(6):e0155677 [FREE Full text] [doi: 10.1371/journal.pone.0155677] [Medline: 27249056]
17. Aylett M, Marples G, Jones K. Home blood pressure monitoring: its effect on the management of hypertension in general practice. Br J Gen Pract 1999 Sep;49(446):725-728 [FREE Full text] [Medline: 10756615]
18. Beevers G, Lip GY, O'Brien E. ABC of hypertension. Blood pressure measurement. Part I-sphygmomanometry: factors common to all techniques. BMJ 2001 Apr 21;322(7292):981-985 [FREE Full text] [doi: 10.1136/bmj.322.7292.981] [Medline: 11312235]
19. Solá J, Proença M, Chételat O. Wearable PWV technologies to measure Blood Pressure: eliminating brachial cuffs. Conf Proc IEEE Eng Med Biol Soc 2013;2013:4098-4101. [doi: 10.1109/EMBC.2013.6610446] [Medline: 24110633]
20. Schoot T, Weenk M, van de Belt TH, Engelen LJLPG, van Goor H, Bredie SJH. A New Cuffless Device for Measuring Blood Pressure: A Real-Life Validation Study. J Med Internet Res 2016 Dec 05;18(5):e85 [FREE Full text] [doi: 10.2196/jmir.5414] [Medline: 27150527]
21. Weenk M, van Goor H, van Acht M, Engelen LJ, van de Belt TH, Bredie SJH. A smart all-in-one device to measure vital signs in admitted patients. PLoS One 2018;13(2):e0190138 [FREE Full text] [doi: 10.1371/journal.pone.0190138] [Medline: 29432461]
22. Gelbart B, Hiscock R, Barfield C. Assessment of neonatal resuscitation performance using video recording in a perinatal centre. J Paediatr Child Health 2010 Jul;46(7-8):378-383. [doi: 10.1111/j.1440-1754.2010.01747.x] [Medline: 20546105]
23. van der Heide PA, van Toledo-Eppinga L, van der Heide M, van der Lee JH. Assessment of neonatal resuscitation skills: a reliable and valid scoring system. Resuscitation 2006 Nov;71(2):212-221. [doi: 10.1016/j.resuscitation.2006.04.009] [Medline: 16987590]
24. Slade E, Williams M, Dwivedi Y. An extension of the UTAUT 2 in a healthcare context. 2013 Mar 19 Presented at: UK Academy for Information Systems Conference Proceedings; March 19, 2013; London p. 55 URL: https://aisel.aisnet.org/ukais2013/55
25. Venkatesh V, Morris M, Davis G, Davis F. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
26. Pluijlaar I. Differences between users and their user problems: a usability test for Wehkamp. 2011 Nov. URL: https://essay.utwente.nl/61686/1/Pluijlaar,_I._-_s0160857_(verslag).pdf [accessed 2019-06-04] [WebCite Cache ID 78ssm8UmB]
27. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation in Industry. London: Taylor and Francis; 1996:1-22.
28. Bangor A. Determining what individual SUS scores mean: adding an adjective rating scale. Journal of Usability Studies 2009;4(3):114-123.
29. Poon C, Zhang YT. Cuff-less and noninvasive measurements of arterial blood pressure by pulse transit time. Conf Proc IEEE Eng Med Biol Soc 2005;6:5877-5880. [doi: 10.1109/IEMBS.2005.1615827] [Medline: 17281597]
30. Boubouchairopoulou N, Kollias A, Chiu B, Chen B, Lagou S, Anestis P, et al. A novel cuffless device for self-measurement of blood pressure: concept, performance and clinical validation. J Hum Hypertens 2017 Dec;31(7):479-482. [doi: 10.1038/jhh.2016.101] [Medline: 28124684]
31. Powers B, Olsen MK, Smith VA, Woolson RF, Bosworth HB, Oddone EZ. Measuring blood pressure for decision making and quality reporting: where and how many measures? Ann Intern Med 2011 Jun 21;154(12):781-8, W. [doi: 10.7326/0003-4819-154-12-201106210-00005] [Medline: 21690592]
32. Pretto J, Roebuck T, Beckert L, Hamilton G. Clinical use of pulse oximetry: official guidelines from the Thoracic Society of Australia and New Zealand. Respirology 2014 Jan;19(1):38-46 [FREE Full text] [doi: 10.1111/resp.12204] [Medline: 24251722]

XSL•FO
RenderX

33. Nitzan M, Romem A, Koppel R. Pulse oximetry: fundamentals and technology update. Med Devices (Auckl) 2014;7:231-239 [FREE Full text] [doi: 10.2147/MDER.S47319] [Medline: 25031547]
34. Teng X, Zhang YT. Theoretical study on the effect of sensor contact force on pulse transit time. IEEE Trans Biomed Eng 2007 Aug;54(8):1490-1498. [doi: 10.1109/TBME.2007.900815] [Medline: 17694870]
35. O'Brien E, Atkins N, Stergiou G, Karpettas N, Parati G, Asmar R, Working Group on Blood Pressure Monitoring of the European Society of Hypertension. European Society of Hypertension International Protocol revision 2010 for the validation of blood pressure measuring devices in adults. Blood Press Monit 2010 Feb;15(1):23-38. [doi: 10.1097/MBP.0b013e3283360e98] [Medline: 20110786]
36. Jeck T, Edmonds D, Mengden T, Schubert M, Vetter W. [Performing self-measurement of blood pressure: a patient survey]. Schweiz Rundsch Med Prax 1991 Apr 23;80(17):456-461. [Medline: 1871472]
37. Lindahl C, Wagner S, Uldbjerg N, Schlütter JM, Bertelsen O, Sandager P. Effects of context-aware patient guidance on blood pressure self-measurement adherence levels. Health Informatics J 2019 Jun;25(2):417-428. [doi: 10.1177/1460458217717073] [Medline: 28701078]
38. Milot J, Birnbaum L, Larochelle P, Wistaff R, Laskine M, Van Nguyen P, et al. Unreliability of home blood pressure measurement and the effect of a patient-oriented intervention. Can J Cardiol 2015 May;31(5):658-663. [doi: 10.1016/j.cjca.2015.03.006] [Medline: 25936491]
39. Wagner S, Kamper CH, Toftegaard TS, Bertelsen OW. Blood pressure self-measurement in the obstetric waiting room. Telemed J E Health 2013 Nov;19(11):872-874. [doi: 10.1089/tmj.2013.0065] [Medline: 24093915]
40. Mengden T, Chamontin B, Phong Chau N, Luis Palma Gamiz J, Chanudet X. User procedure for self-measurement of blood pressure. First International Consensus Conference on Self Blood Pressure Measurement. Blood Press Monit 2000 Apr;5(2):111-129. [Medline: 10828898]
41. Bonderup A, Hangaard SV, Lilholt PH, Johansen MD, Hejlesen OK. Patient support ICT tool for hypertension monitoring. Stud Health Technol Inform 2012;180:189-193. [Medline: 22874178]
42. Strandberg TE, Pitkala K. What is the most important component of blood pressure: systolic, diastolic or pulse pressure? Curr Opin Nephrol Hypertens 2003 May;12(3):293-297. [doi: 10.1097/01.mnh.0000069868.94246.ef] [Medline: 12698068]
43. Korhonen I, Pärkkä J, van Gils M. Health monitoring in the home of the future. IEEE Eng Med Biol Mag 2003;22(3):66-73. [Medline: 12845821]
44. Mengden T, Hernandez Medina RM, Beltran B, Alvarez E, Kraft K, Vetter H. Reliability of reporting self-measured blood pressure values by hypertensive patients. Am J Hypertens 1998 Dec;11(12):1413-1417. [doi: 10.1016/s0895-7061(98)00241-6] [Medline: 9880121]
45. Vasileiou K, Barnett J, Young T. Interpreting and acting upon home blood pressure readings: a qualitative study. BMC Fam Pract 2013 Jul 13;14:97 [FREE Full text] [doi: 10.1186/1471-2296-14-97] [Medline: 23849055]
46. Rickerby J, Woodward J. Patients' experiences and opinions of home blood pressure measurement. J Hum Hypertens 2003 Jul;17(7):495-503. [doi: 10.1038/sj.jhh.1001582] [Medline: 12821957]
47. Mancia G, Fagard R, Narkiewicz K, Redon J, Zanchetti A, Böhm M, Task Force for the Management of Arterial Hypertension of the European Society of Hypertensionthe European Society of Cardiology. 2013 ESH/ESC Practice Guidelines for the Management of Arterial Hypertension. Blood Press 2014 Feb;23(1):3-16. [doi: 10.3109/08037051.2014.868629] [Medline: 24359485]
48. Lewis J, Sauro R. The Factor Structure of the System Usability Scale. 2009 Presented at: First International Conference, HCD 2009, Held as Part of HCI International 2009; July 19-24, 2009; San Diego, CA, USA p. 94-103. [doi: 10.1007/978-3-642-02806-9_12]
49. Zhou X, Wu W, Bao S. [Mobile Health: IEEE Standard for Wearable Cuffless Blood Pressure Measuring Devices]. Zhongguo Yi Liao Qi Xie Za Zhi 2015 Jul;39(4):285-287. [Medline: 26665952]

Abbreviations

BP: blood pressure
CE: Conformité Européene
DBP: diastolic blood pressure
ECG: electrocardiogram
FDA: Food and Drug Administration
HBPM: home blood pressure monitoring
SBP: systolic blood pressure
SpO_2: peripheral capillary oxygen saturation
SUS: System Usability Scale
UTAUT2: Unified Theory of Acceptance and Use of Technology 2

XSL•FO
RenderX

Edited by G Eysenbach; submitted 31.05.18; peer-reviewed by Y Yaniv, C Exeter, S Kara, B Green, M Lang; comments to author 06.10.18; revised version received 28.02.19; accepted 29.04.19; published 19.06.19.

Please cite as:
Ogink PAM, de Jong JM, Koeneman M, Weenk M, Engelen LJLPG, van Goor H, van de Belt TH, Bredie SJH
Feasibility of a New Cuffless Device for Ambulatory Blood Pressure Measurement in Patients With Hypertension: Mixed Methods Study
J Med Internet Res 2019;21(6):e11164
URL: http://www.jmir.org/2019/6/e11164/
doi:10.2196/11164
PMID:31219050

©Paula AM Ogink, Jelske M de Jong, Mats Koeneman, Mariska Weenk, Lucien JLPG Engelen, Harry van Goor, Tom H van de Belt, Sebastian JH Bredie. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Feasibility of a Real-Time Clinical Augmented Reality and Artificial Intelligence Framework for Pain Detection and Localization From the Brain

Xiao-Su Hu[1,2], MSc, PhD; Thiago D Nascimento[1], DDS, MS; Mary C Bender[1], BS; Theodore Hall[3], PhD; Sean Petty[3], BS; Stephanie O'Malley[3], BFA; Roger P Ellwood[4], PhD; Niko Kaciroti[1,2,5], PhD; Eric Maslowski[6], BS; Alexandre F DaSilva[1,2], DDS, DMedSc

[1]Headache & Orofacial Pain Effort Lab, Biologic & Materials Sciences Department, School of Dentistry, University of Michigan, Ann Arbor, MI, United States

[2]Center for Human Growth and Development, University of Michigan, Ann Arbor, MI, United States

[3]3D Lab, Digital Media Commons, University of Michigan, Ann Arbor, MI, United States

[4]Clinical Method Development, Colgate Palmolive, Piscataway, NJ, United States

[5]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[6]Moxytech Inc, Ann Arbor, MI, United States

Corresponding Author:
Alexandre F DaSilva, DDS, DMedSc
Headache & Orofacial Pain Effort Lab, Biologic & Materials Sciences Department
School of Dentistry
University of Michigan
1011 N University Ave
Ann Arbor, MI, 48109-1078
United States
Phone: 1 734 615 3807
Email: adasilva@umich.edu

Abstract

Background: For many years, clinicians have been seeking for objective pain assessment solutions via neuroimaging techniques, focusing on the brain to detect human pain. Unfortunately, most of those techniques are not applicable in the clinical environment or lack accuracy.

Objective: This study aimed to test the feasibility of a mobile neuroimaging-based clinical augmented reality (AR) and artificial intelligence (AI) framework, CLARAi, for objective pain detection and also localization direct from the patient's brain in real time.

Methods: Clinical dental pain was triggered in 21 patients by hypersensitive tooth stimulation with 20 consecutive descending cold stimulations (32°C-0°C). We used a portable optical neuroimaging technology, functional near-infrared spectroscopy, to gauge their cortical activity during evoked acute clinical pain. The data were decoded using a neural network (NN)–based AI algorithm to classify hemodynamic response data into pain and no-pain brain states in real time. We tested the performance of several networks (NN with 7 layers, 6 layers, 5 layers, 3 layers, recurrent NN, and long short-term memory network) upon reorganized data features on pain diction and localization in a simulated real-time environment. In addition, we also tested the feasibility of transmitting the neuroimaging data to an AR device, HoloLens, in the same simulated environment, allowing visualization of the ongoing cortical activity on a 3-dimensional brain template virtually plotted on the patients' head during clinical consult.

Results: The artificial neutral network (3-layer NN) achieved an optimal classification accuracy at 80.37% (126,000/156,680) for pain and no pain discrimination, with positive likelihood ratio (PLR) at 2.35. We further explored a 3-class localization task of left/right side pain and no-pain states, and convolutional NN-6 (6-layer NN) achieved highest classification accuracy at 74.23% (1040/1401) with PLR at 2.02.

Conclusions: Additional studies are needed to optimize and validate our prototype CLARAi framework for other pains and neurologic disorders. However, we presented an innovative and feasible neuroimaging-based AR/AI concept that can potentially

XSL•FO
RenderX

transform the human brain into an objective target to visualize and precisely measure and localize pain in real time where it is most needed: in the doctor's office.

International Registered Report Identifier (IRRID): RR1-10.2196/13594

(J Med Internet Res 2019;21(6):e13594) doi:10.2196/13594

KEYWORDS

pain; spectroscopy, near-infrared; virtual reality; artificial intelligence

Introduction

Background

Accurate pain assessment is crucial across a wide range of acute and chronic pain conditions to provide proper diagnosis and treatment, especially when patients have limitations to express their ongoing suffering. The estimated economic impact of pain, from direct medical costs to loss of productive time, is US $560 to $635 billion every year [1]. Despite this, we are still heavily relying on the following measure: "From 0 to 10—0 being no pain and 10 being the worst pain—what is your pain level?" All of our clinical and research decisions on the efficacy of current or new potential pain therapies and, most importantly, our patients' ongoing pain levels, are biased by the inaccuracy of that scale, independent of how rigorous, complex, and costly our protocols are.

The pain field has progressed by quantifying the patients' suffering with more holistic pain questionnaires and measure scales (eg, McGill Pain Questionnaire and Face Rating Pain Scale), which are prevalent, useful, and convenient. However, the subjective reports still carry limitations: first, they are inconsistent among different patient groups regarding age and cultures. For instance, words used by patients nowadays to express the severity of their pain have also evolved with time and might be different from the ones articulated by past generations [2]. Second, those tools cannot be applied during procedures or surgeries that impair patients' communication, including the minimally conscious or cognitively impaired. Finally, self-report provides limited value for understanding the neurophysiological processes underlying different types of pain, thereby, blurring the treatments to the underlying neuropathologic conditions [3].

To address these limitations, researchers have started to analyze the neurological signature of pain using neuroimaging [3,4]. Wager and colleagues developed a system using machine learning technology on data collected with functional magnetic resonance imaging (fMRI), showing the possibility of detecting a robust neurological signature of pain at the level of the individual person. Other researchers have demonstrated the possibility of detecting even temporomandibular disorder using multivoxel pattern analysis on fMRI signal. Such MRI-based gold rush to report new brain-pain biomarkers forced the field to recommend standards of evidence [5]. Indeed, fMRI objective assessments of pain have provided a great step ahead in the path of dissecting brain mechanism of pain, but the size and cost of the MRI scanner and other conventional neuroimaging tools (eg, positron emission tomography) prevent its application in the clinical office. This impediment has sparked the interest of portable neuroimaging devices with similar technical benefits as fMRI. Functional near-infrared spectroscopy (fNIRS) detects concentration variations of oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HbR), such as blood oxygen level dependent signal in fMRI. It measures the absorption of near-infrared light at wavelengths between 700 and 1000 nm, noninvasively through the skull [6,7]. Compared with the MRI scanner, the portability and compatibility to ferromagnetic/electrical components provide researchers an option to monitor, localize, and analyze functional brain activity in the surgical and clinical environment [8-12].

Objectives

In previous studies, our group studied the hemodynamic cortical responses detected by fNIRS in patients with hypersensitive teeth in the dental chair [13]. We found well-defined hemodynamic cortical activity in the primary sensory (S1) and prefrontal cortices (PFCs) elicited by thermal stimulation to the affected tooth from expectation to pain detection. Interestingly, the patients' clinical pain experience was predicted concomitantly by their baseline functional connectivity between S1 and PFC, as well as a well-defined stepwise sequence of hemodynamic responses. This sensory-discriminative and cognitive-emotional cascade of brain responses initiated during the expectation of the clinical pain (prepain phase), with activations in the contralateral S1 orofacial homuncular region and also in the bilateral PFC. Such activations were followed by flat or PFC deactivation and further S1 responses when the cold stimuli crossed noxious levels (pain phase) [14]. Herein, following our earlier findings, we aimed to develop an in-house framework technology that can visualize, measure, and decode in real time the ongoing cascade of spread cortical activities into when and where there is clinical pain. This was successfully achieved through 3-step experiments integrating optical neuroimaging (fNIRS), augmented reality (AR), and a neural network (NN)–based artificial intelligence (AI).

Methods

Data Acquisition

The University of Michigan Institutional Review Board approval was obtained before study initiation. We recruited 21 participants (8 male; age: mean 27.6, SD 3.5 years) with hypersensitive teeth. We collected neuroimaging data from a thermal stimulation session [13,14]. In this session, the participants underwent 20 thermal stimulation trials, in which the thermal probe cycled from 32^{o}C to 0^{o}C at a rate of –2^{o}/second. Subjects controlled the cooling unit by clicking a computer mouse when clinical pain was achieved, causing the stimulation to stop.

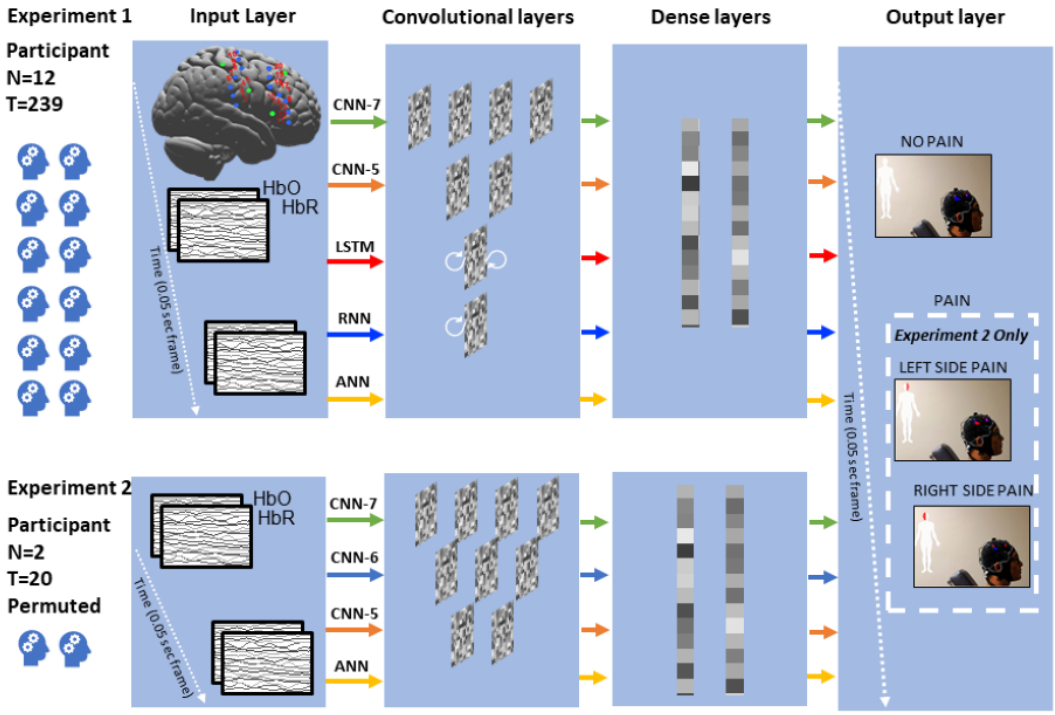
XSL•FO
RenderX

The data were acquired with a TechEN-CW6 fNIRS (Milford, MA, United States) system at a 20 Hz sampling rate. The setup included 8 emitters of near-infrared light and 28 detectors spaced 3 cm apart, yielding 40 data channels deployed at bilateral PFC and S1. The probe set was designed based on the international 10-10 transcranial positioning system [15] and further validated with automatic anatomical labeling database [16]. The collected raw data were examined by a 2-step data quality control steps, filtered with a band-pass filter at 0.01-0.3 Hz and then converted to HbO and HbR concentration change data. Such preprocessing was completed using scripts from Homer2 software (Huppert et al) [17] and several custom MATLAB (MathWorks, MA, United States) scripts.

In this study, we employed 2 experiments for testing the feasibility of pain/no-pain prediction as well as left/right pain localization (Figure 1). In addition, we conducted a third experiment of designing an AR-based data visualization terminal. The entire study framework is presented in Figure 2.

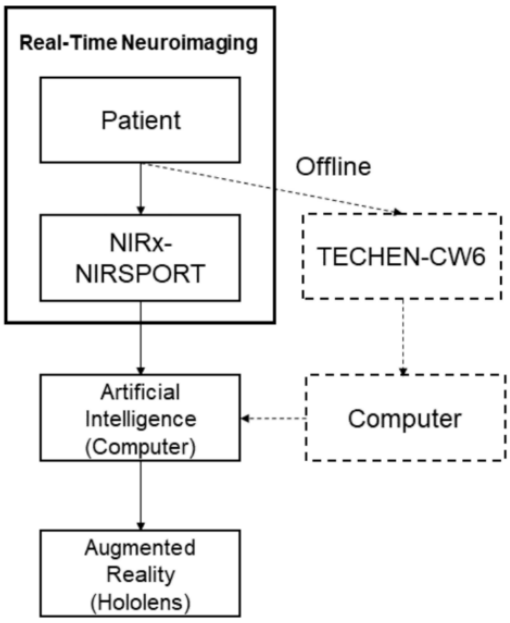
The NN design, training, and testing were completed in a Python-based toolbox, Keras (Chollet et al) [18], with Tensorflow backend (Google Brain Team) [19] and Scikit-learn (Cournapeau et al) [20] toolbox for cross-validation. The data displaying terminal on the AR device were designed through HoloBrain (Microsoft, WA, United States), an in-house developed software at University of Michigan.

Figure 1. Experiment flow chart. The green line indicates the convolutional neural network with 7 layers (CNN-7), the blue line indicates the CNN network with 6 layers, the orange line indicates the CNN network with 5 layers (CNN-5), the red line indicates the long short-term memory network, the dark blue line indicates the recurrent NN, and the yellow line indicates the artificial NN with 3 layers for experiment 1—pain/no-pain prediction and experiment 2—left/right pain localization task. Experiment 1 included the data collected from N=12 participants, 239 trials in total, whereas experiment 2 included the data collected from N=2 participants, 20 trials in total. CNN: convolutional neural network.



XSL•FO
RenderX

Figure 2. Study framework.



Experiment 1: Pain Detection

The aim of the experiment was to test the feasibility of pain/no-pain prediction at individual patient level. We tested convolutional NN (CNN) configurations at 3 different depths, respectively, 7 layers (CNN-7), 5 layers (CNN-5), and 3 layers (artificial NN, ANN) to evaluate their performance on same datasets (Figure 1). In addition, we also tested a recurrent NN and a long short-term memory network on our dataset considering the possible temporal connection within fNIRS time series. Prior studies suggested that including data history as feature can improve classification performance [9,21]. In our most recent research, we reported the cascade of brain events during clinical pain that demonstrated interactive pain expectation evoked responses at bilateral prefrontal cortices as well as a 2-peak response at contralateral sensory cortex [13]. Using that knowledge in this study, we assembled 2 types of feature for the input layer: (1) a 40 × 40 × 2 data cube, by including 40 samples (2-second data history block sampled at 20 Hz) and 40 channels with 2 types of data (HbO/HbR) in the third dimension and (2) a 80 × 40 × 2 data cube, by including 80 samples (2 × 2-second data history blocks before and after patients' pain threshold). Given that there were more no-pain than pain samples in the dataset within the multiple clinical stimulation trials in each individual, we balanced the 2 sample sets by reweighting their loss functions during the training process [22]. Specifically, we defined a dictionary for the pain/no-pain labels with associated weights of 10:1 and assigned such weight during training in the Keras toolbox using a class_weight variable. In addition, we used 10-fold cross-validation to validate each model [23] and calculated the averaged classification accuracy, sensitivity, specificity, positive likelihood ratio (PLR), positive predictive value, negative predictive value, and kappa value to evaluate the classification performance.

Experiment 2: Pain Localization

The aim of the experiment was to further test the feasibility of left/right pain and no-pain states prediction (3-class classification) on merged and permuted patients' data (data were collected from patient 3 and 19, separately, left/right tooth stimulated). We permuted the merged data by randomly including and excluding data cubes along time course. We tested all networks applied in experiment 1, and in addition a 6-layer CNN (CNN-6) on type I data cube, with time series preprocessed with a custom real-time normalization algorithm (Figure 1). Specifically, we used a divided by the mean [24] scheme to normalize the data in real time, where the mean was updated by a windowed data along the time course. To retain the uniformity in the data, we did not run cross-validation in this scenario.

Experiment 3: Augmented Reality—Pain Visualization and Decoding Using Augmented Reality and Artificial Intelligence

We developed a displaying terminal for the framework using an AR device, HoloLens (Microsoft, WA, United States). AR is a computer vision–based technology that expands our real world by adding a layer of virtual and digital information to it. It is becoming prevalent in different fields including, for example, construction, gaming, and medicine. The Hololens is a headset-shaped AR computer developed by Microsoft, which allows users to visualize 3-dimensional (3D) holographic images on top of the real physical world. In this study, the functional hemodynamic response data acquired from the patient's brain at multiple cortical regions of interest were wirelessly transmitted to the HoloLens device. Afterwards, we used an in-house–developed software to display the ongoing patient's cortical function updating in real time on the brain template modeled in the software (Multimedia Appendix 1; video section 1 and 2). This software is an application adapted from an in-house 3D rendering engine developed at the University of Michigan 3D Lab. In active development for over 10 years, this platform allows for the rapid development and displaying of

XSL•FO
RenderX

complex interactive 3D scenes including advanced materials, lighting, physical responses, and detailed meshes. In this study, we first registered the 40 data acquisition channels to an MNI 152 nonlinear brain template. Afterwards, we reconstructed this virtual brain with the registered functional regions using this software to display and adjust its appearance based on incoming data from the NIRS device. Furthermore, the software decoded and displayed the brain activity from experiments 1 and 2 in clinical pain/no-pain status and localization by mapping the ongoing results on a virtually reconstructed digital body within the field of AR view (Multimedia Appendix 1; video section 2). In addition, the CLARAi displaying of ongoing cortical activity in volunteers was also tested in real time using a NIRSport fNIRS system with 16 source-detector density (NIRx, NY, United States) to facilitate the use of our CLARAi concept in real clinical environment (Multimedia Appendix 1; video section 1).

Results

Experiment 1 and 2

Of the 21 participants, 12 were further preprocessed to enter the feasibility testing in experiment 1 (Table 1). Within these 12 participants, we had a total 180,580 data cubes; among these cubes, 23,900 (13.24%, 23,900/180,580) were labeled as pain and 156,680 (86.76%, 156,680/180,580) were labeled as no-pain. A total of 2 participants' data were tested in experiment 2; there were 30,820 data cubes in total, with 2000 (6.49%, 2000/30,820) labeled as right-side pain, 2000 (6.49%, 2000/30,820) labeled as left side pain, and 26,820 (87.02%, 26,820/30,820) labeled as no pain. Figure 3 shows representative averaged HbO and HbR responses across all channels under pain and no-pain statues, respectively, collected from patient 3.

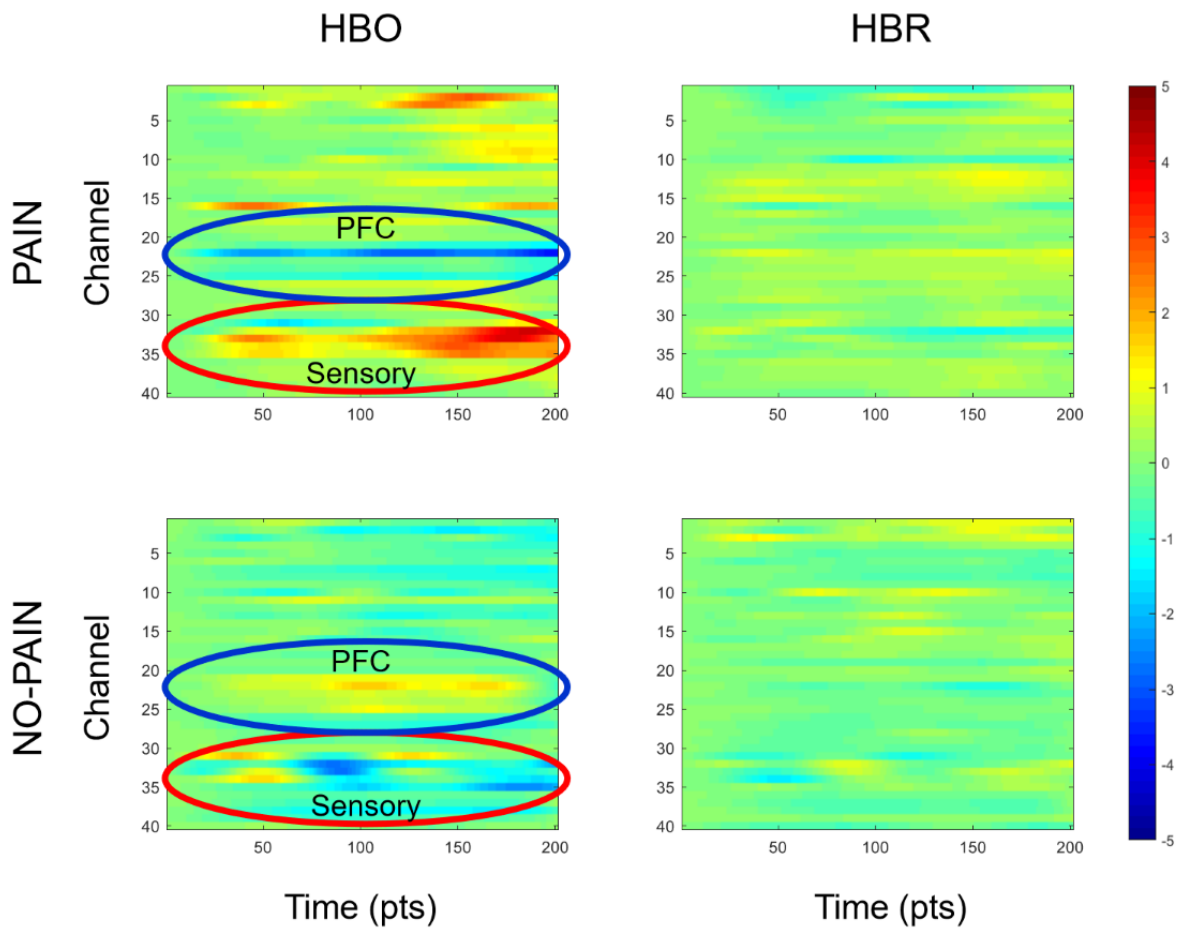
Table 1. Participant demographics with classification performance.

Participant	Pain (points)	No-pain (points)	Class accuracy	Reported NRS[a]	Stimulation side
3	2000	14,980	84.58 (%)	5.5	Right
5	2000	13,080	76.83 (%)	3.9	Right
10	2000	12,140	78.84 (%)	5.8	Right
11	2000	11,520	80.98 (%)	1.9	Left
12	2000	13,320	81.53 (%)	3.4	Left
13	2000	13,700	76.25 (%)	8.5	Right
15	2000	11,480	76.12 (%)	5.8	Left
16	2000	12,600	79.55 (%)	6.6	Left
17	1900	14,360	80.74 (%)	2.6	Left
18	2000	13,020	82.29 (%)	3.8	Left
19	2000	11,840	81.00 (%)	5.1	Left
20	2000	14,640	85.78 (%)	3.3	Left

[a]NRS: numerical rating scale.

XSL•FO
RenderX

Figure 3. Representative averaged oxygenated hemoglobin (HbO) and deoxygenated hemoglobin (HbR) heat map from all data channels. The upper and lower panels, respectively, indicated the hemodynamic responses during pain and no-pain statues. The left and right panels, respectively, indicated the HbO and HbR responses. The red and blue circles, respectively, highlighted 2 regions of interest, sensory and prefrontal cortices. HbO: oxygenated hemoglobin; HbR: deoxygenated hemoglobin; PFC: prefrontal cortex.



XSL•FO
RenderX

Table 2. Performance of different network setups in experiment 1.

Network setup	Overall accuracy	Sensitivity	Specificity	PPV[a]	NPV[b]	PLR[c]	Kappa
CNN[d]-7	79.62 (%)	0.144	0.896	0.169	0.872	1.39	0.04
CNN-5	79.25 (%)	0.153	0.891	0.183	0.872	1.4	0.05
ANN[e]	79.17 (%)	0.192	0.884	0.205	0.877	1.65	0.08
ANN+2 portion	80.37 (%)	0.326	0.861	0.266	0.893	2.35	0.17
ANN+2 portion + oversample	75.93 (%)	0.409	0.801	0.242	0.898	2.06	0.16
ANN+2 portion + oversample (HbO[f] only)	77.19 (%)	0.379	0.819	0.245	0.895	2.10	0.16
RNN[g]+2 portion + oversample	76.31 (%)	0.332	0.815	0.211	0.888	1.80	0.11
LSTM[h]+2 portion + oversample	77.29 (%)	0.319	0.828	0.220	0.887	1.86	0.12

[a]PPV: positive predictive value.
[b]NPV: negative predictive value.
[c]PLR: positive likelihood ratio.
[d]CNN: convolutional neural network.
[e]ANN: artificial neural network.
[f]HbO: oxygenated hemoglobin.
[g]RNN: recurrent neural network.
[h]LSTM: long short-term memory.

Table 3. Performance of different network setups in experiment 2.

Network	Accuracy	Sensitivity	Specificity	PPV[a]	NPV[b]	PLR[c]	Kappa
ANN[d]	70.88 (%)	0.443	0.777	0.339	0.844	1.99	0.20
CNN[e]-5	65.37 (%)	0.375	0.723	0.250	0.824	1.35	0.08
CNN-6	74.23 (%)	0.279	0.862	0.342	0.823	2.02	0.15
CNN-7	73.23 (%)	0.540	0.782	0.389	0.868	2.48	0.28

[a]PPV: positive predictive value.
[b]NPV: negative predictive value.
[c]PLR: positive likelihood ratio.
[d]ANN: artificial neural network.
[e]CNN: convolutional neural network.

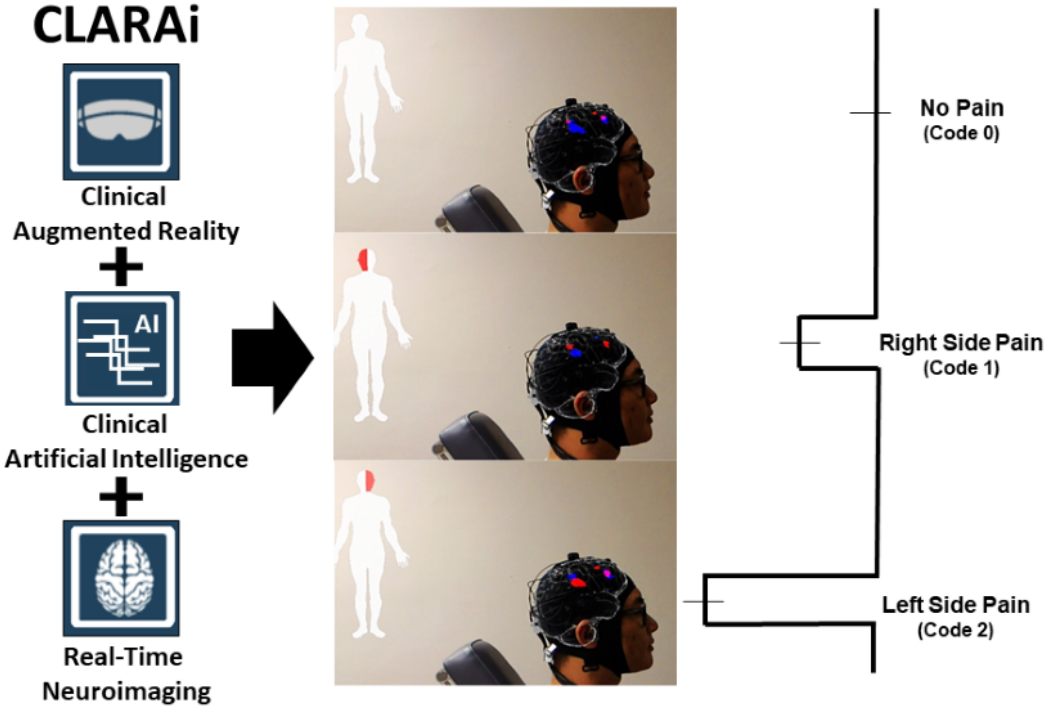
ANN performed on split data history segments achieved the best results, with a prediction accuracy at 80.37 % 145,210/180,580), and a PLR at 2.35 (sensitivity=0.326, specificity=0.861). ANN performed on split data history blocks with reweighted loss function achieved the highest sensitivity at 0.409 and specificity of 0.801, with a PLR at 2.06. In addition, CNN-7 achieved the highest specificity at 0.896, however, with a PLR at 1.39 and a sensitivity of 0.144. A detailed performance summary of experiment 1 for different network setups can be found in Table 2. In addition, a patient-wise classification accuracy was listed in Table 1. The 3-class prediction achieved an optimal classification accuracy at 74.23% (1040/1401) with CNN-6, with a PLR at 2.02 in real time, whereas CNN-7 had a highest sensitivity at 0.540 with a PLR at 2.48. A detailed performance summary of experiment 2 can be found in Table 3.

Experiment 3

Figure 4 shows the developed data displaying interface "HoloBrain." The collected functional HbO and HbR data were displayed and updated on an MNI152 brain template in real time (Multimedia Appendix 1; video section 1). The 3D virtual brain activation image, through HoloLens, was superimposed onto a participant's head. Beside the 3D brain activation, an animated human body with modulated red areas were indicating pain regions prediction by side—either left or right cranio-orofacial regions (Multimedia Appendix 1; video section 2).

XSL•FO
RenderX

Figure 4. CLARAi framework that integrated clinical real-time neuroimaging, augmented reality, and artificial intelligence provides an augmented clinical environment by displaying neuroimaging data with predicted and localized pain of patient. The classification codes for no-pain, right side pain, and left side pain was defined as 0, 1, and 2, respectively, for model training purposes.



Discussion

In our previous study, we observed, respectively, clinical pain expectation and pain-related responses at PFC and S1 cortices [13]. Further in our successive report, we discovered the sequential connections among those cortical responses, meaning the cascade of cortical events in the brain before, during, and after the clinical pain experience [14]. In this feasibility study, we assembled that pattern for data signature for objective clinical pain prediction using fNIRS data collected directly from the bilateral PFC and S1 cortices. Previous studies achieved promising results in attempting to classify different level of thermal stimulation (potentially indicating pain vs no-pain) based on hemodynamic response data collected from sensory cortex [25,26]. These studies also examined the performance of several prevalent machine learning methods including support vector machine, linear discriminant analysis, and K-nearest neighbor. In our study, we chose to examine the performance of different NN setups, given our data were collected from multiple regions of interest bilaterally including PFCs and S1 cortices at a relatively high sampling rate. On the basis of our trilogy of experiments, our CLARAi model gained not only information from spatial pattern but also temporal sequence in the data by including up to 10-second data history counting back from each data frame to get the contrast between clinical pain expectation and pain experience per se in real time.

Herein, in experiment 1, we tested several NNs on different reorganized brain activation data to predict pain and no-pain conditions. We first tested 3 networks on data including 2-second data history block and found that CNN-7 achieved the highest general classification accuracy. In recent years, CNNs became deeper and deeper, with state-of-the-art networks going from 7 layers AlexNet [27] to a thousand layers Residual Nets [28] in a 4-year period. The reason behind the boost is that a deeper network can usually learn a more complex nonlinear function. Our results on type I data cube complied with such trend that CNN-7 achieved slightly higher accuracy. However, the CNN-7 results emphasized much more specificity than sensitivity, meaning we were likely to better define if the patient had no-pain than pain. Such results may be because of several potential reasons: (1) The collected data were unbalanced, meaning we had much more no-pain than pain samples in the training set from each patient; (2) We did not have enough data to train the large number of parameters for CNN; and (3) The pattern contained within the 2-second data history may not be informative enough for discriminating pain from no-pain condition. Therefore, to improve sensitivity, we employed a simpler ANN network, reweighted the loss function relevant to the pain condition samples, and assembled data feature with 2 split history blocks to include pain expectation evoked responses [29]. When the preceding pain expectation (prepain) phase is also taken into consideration in the algorithm [29], the results indicated that the real time sensitivity was significantly increased to detect pain, from 0.144 to 0.409, while not losing too much specificity, from 0.896 to 0.861. Such improvement in pain detection (sensitivity) suggested that the pain expectation phase was crucial to encode the magnitude of the immediate pain and

XSL•FO
RenderX

is highly driven by the activation of the left dorsolateral PFC [14,30]. In addition, we tested the same NN setup on only HbO data. We found a slight increase in general classification accuracy and specificity (0.013 and 0.018), whereas a decrease in sensitivity (0.029). As sensitivity was generally lower than specificity in this study, we selected combined HbO and HbR data as features for classification. Finally, we examined whether there was a potential correlation between the individual level classification accuracy and the reported numerical rating scale of pain, but we did not find any statistical significance.

Considering the relatively high-spatial resolution of fNIRS imaging, in experiment 2, we further tested the feasibility of localizing pain. We introduced a 3-class pain localization problem by merging the data from 2 selected patients, one with left side tooth pain hypersensitivity and the other one with same clinical condition on a right tooth during cold stimulation. To eliminate baseline and signal magnitude difference, we applied a simulated real-time normalization algorithm to the data. We then tested this dataset with several NNs with different depth and found that CNN-6 achieved the best general classification accuracy. Though there is need for further validation, the preliminary discrimination result demonstrated a strong potential of our framework in localizing pain at different body regions. Moreover, the results demonstrated the feasibility of training a universal model that can localize pain condition across patients based on the S1 homuncular activation by side and major body regions. This is biologically feasible because the somatotopic homuncular S1 representation for pain in the orofacial region is quite large, like other major functional body regions including the thumb/hand, trunk, and feet [31].

Combined with the pain prediction module, we developed a clinical AR-based data displaying interface for the framework. The data collected in this study with the predicted result were transferred to a HoloLens device. The magnitudes of hemodynamic response changes at multiple locations on a 3D brain template were superimposed on the participants' head in reality via HoloLens (Multimedia Appendix 1; video 1). The predicted painful locations were indicated by the red flickering parts on the animated body beside the virtual brain. In a true clinical environment, with such framework, clinicians can better understand in an objective way to determine when/where the patients are suffering from pain, especially when they cannot express themselves. In addition, the potential idea is to even decide the level of pain, and further a "prepain" phase using PFC activation evoked by pain-associated anxiety or expectation. Such information will help clinicians decide when to intervene for addressing the pain or the immediate likelihood of it to occur (eg, anesthesia and brain stimulation). Afterwards, the entire framework becomes closed loop by including a pain intervention module, for instance pain neuromodulation.

Finally, all selected data preprocessing, classification, transmitting, and displaying methods in this study can be implemented in real time. However, the CLARAi framework is in its initial stages. Future improvements of this work include: (1) optimizing the framework sensitivity by potentially adding short-separation channels during data acquisition to model interfering physiological signals in a better way, (2) expansion of the current participant-specific model to a general model with learning ability that will only require individualization to precisely adapt to variations in each patient, and (3) further expansion of the model to fit other types of pain conditions and neurologic disorders including depression and anxiety. In summary, we tested the feasibility of a prototype of a mobile neuroimaging-based clinical AR and AI (CLARAi) framework for objective pain detection and localization in the clinical environment in real time. Such framework predicted when and where there was physical pain based on the brain statuses in our data study and displayed neuroimaging data interactively in real time. Although extensive validation work still needs to be done, the CLARAi framework might turn into reality the goal of precisely "seeing and believing" the biologic pain suffering of our patients in the doctor's office.

Acknowledgments

This research was sponsored by an unrestricted grant from the Colgate Palmolive Company and fNIRS pilot grant award (University of Michigan, UM). Drs Hu, DaSilva, Hall, Petty, Maslowski are the co-creators of CLARAi. We would also like to thank Andrew Racek for his contribution to data collection.

Conflicts of Interest

The content described within this study has been developed at the UM and disclosed to the UM Office of Technology Transfer. All intellectual property rights including but not limited to patents/patent applications, trademark and copyright of software, algorithms, reports, displays, and visualizations are owned by the Regents of the University of Michigan. Drs DaSilva and Maslowski are the co-creators of CLARAi. Drs DaSilva and Malowski are co-founders of MoxyTech Inc, which has optioned the technology CLARAi from University of Michigan. Dr. Roger P Ellwood, who was previously an employee of the Colgate Palmolive company-the funding agency of this work.

Multimedia Appendix 1

Video: section 1, the collected oxygenated hemoglobin and deoxygenated hemoglobin data were displayed on an MNI152 brain template in real time. Section 2, the 3-dimensional (3D) virtual brain activation image, through HoloLens, was superimposed onto a participant's head. Beside the 3D brain activation, an animated human body with modulating red areas indicated pain regions prediction by side—either left or right cranio-orofacial regions.

XSL•FO
RenderX

[MP4 File (MP4 Video), 141MB - jmir_v21i6e13594_app1.mp4]

References

1. Gaskin DJ, Richard P. The economic costs of pain in the United States. J Pain 2012 Aug;13(8):715-724. [doi: 10.1016/j.jpain.2012.03.009] [Medline: 22607834]
2. Nascimento TD, DosSantos MF, Danciu T, DeBoer M, van Holsbeeck H, Lucas SR, UMSoD (Under)Graduate Class Of 2014, et al. Real-time sharing and expression of migraine headache suffering on Twitter: a cross-sectional infodemiology study. J Med Internet Res 2014 Apr;16(4):e96 [FREE Full text] [doi: 10.2196/jmir.3265] [Medline: 24698747]
3. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo C, Kross E. An fMRI-based neurologic signature of physical pain. N Engl J Med 2013 Apr 11;368(15):1388-1397 [FREE Full text] [doi: 10.1056/NEJMoa1204471] [Medline: 23574118]
4. Harper DE, Shah Y, Ichesco E, Gerstner GE, Peltier SJ. Multivariate classification of pain-evoked brain activity in temporomandibular disorder. Pain Rep 2016 Sep;1(3) [FREE Full text] [doi: 10.1097/PR9.0000000000000572] [Medline: 28630949]
5. Davis KD, Flor H, Greely HT, Iannetti GD, Mackey S, Ploner M, et al. Brain imaging tests for chronic pain: medical, legal and ethical issues and recommendations. Nat Rev Neurol 2017 Oct;13(10):624-638. [doi: 10.1038/nrneurol.2017.122] [Medline: 28884750]
6. Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. Neuroimage 2012 Nov 01;63(2):921-935. [doi: 10.1016/j.neuroimage.2012.03.049] [Medline: 22510258]
7. Obrig H, Villringer A. Beyond the visible--imaging the human brain with light. J Cereb Blood Flow Metab 2003 Jan;23(1):1-18. [doi: 10.1097/01.WCB.0000043472.45775.29] [Medline: 12500086]
8. Naseer N, Hong K. fNIRS-based brain-computer interfaces: a review. Front Hum Neurosci 2015;9:3 [FREE Full text] [doi: 10.3389/fnhum.2015.00003] [Medline: 25674060]
9. Hu X, Hong K, Ge SS. fNIRS-based online deception decoding. J Neural Eng 2012 Apr;9(2):026012. [doi: 10.1088/1741-2560/9/2/026012] [Medline: 22337819]
10. Hu X, Hong K, Ge SS, Jeong M. Kalman estimator- and general linear model-based on-line brain activation mapping by near-infrared spectroscopy. Biomed Eng Online 2010 Dec 08;9:82 [FREE Full text] [doi: 10.1186/1475-925X-9-82] [Medline: 21138595]
11. Yücel MA, Aasted CM, Petkov MP, Borsook D, Boas DA, Becerra L. Specificity of hemodynamic brain responses to painful stimuli: a functional near-infrared spectroscopy study. Sci Rep 2015 Mar 30;5:9469 [FREE Full text] [doi: 10.1038/srep09469] [Medline: 25820289]
12. Aasted CM, Yücel MA, Steele SC, Peng K, Boas DA, Becerra L, et al. Frontal lobe hemodynamic responses to painful stimulation: a potential brain marker of nociception. PLoS One 2016;11(11):e0165226 [FREE Full text] [doi: 10.1371/journal.pone.0165226] [Medline: 27806119]
13. Racek AJ, Hu X, Nascimento TD, Bender MC, Khatib L, Chiego D, et al. Different brain responses to pain and its expectation in the dental chair. J Dent Res 2015 Jul;94(7):998-1003. [doi: 10.1177/0022034515581642] [Medline: 25904140]
14. Hu X, Racek AJ, Bellile E, Nascimento TD, Bender MC, Toback RL, et al. Brain functional changes before, during, and after clinical pain. J Dent Res 2018 May;97(5):523-529. [doi: 10.1177/0022034517750136] [Medline: 29324076]
15. Jurcak V, Tsuzuki D, Dan I. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. Neuroimage 2007 Feb 15;34(4):1600-1611. [doi: 10.1016/j.neuroimage.2006.09.024] [Medline: 17207640]
16. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 2002 Jan;15(1):273-289. [doi: 10.1006/nimg.2001.0978] [Medline: 11771995]
17. Huppert TJ, Diamond SG, Franceschini MA, Boas DA. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. Appl Opt 2009 Apr 01;48(10):D280-D298 [FREE Full text] [Medline: 19340120]
18. Chollet François. Keras. 2015. URL: https://keras.io/ [accessed 2019-05-30] [WebCite Cache ID 78lJPUeAn]
19. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. ArXiv e-prints Internet. 2016. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems URL: https://arxiv.org/abs/1603.04467 [accessed 2019-06-05] [WebCite Cache ID 78uhfWhlw]
20. Fabian PG. Scikit-learn: Machine Learning in Python. J Mach Learn Res 2011;12:2825-2830.
21. Cui X, Bray S, Reiss AL. Speeded near infrared spectroscopy (NIRS) response detection. PLoS One 2010 Nov 11;5(11):e15474. [doi: 10.1371/journal.pone.0015474] [Medline: 21085607]
22. Sun Y, Wong AKC, Kamel MS. Classification of imbalanced data: A review. Int. J. Patt. Recogn. Artif. Intell 2009 Jun;23(04):687-719. [doi: 10.1142/S0218001409007326]
23. Stone M. Cross-validatory choice and assessment of statistical predictions. J R Stat Soc Series B Stat Methodol 1974;36(2):111-147.
24. Reades J, Calabrese F, Sevtsuk A, Ratti C. Cellular census: explorations in urban data collection. IEEE Pervasive Comput 2007 Jul;6(3):30-38. [doi: 10.1109/MPRV.2007.53]

XSL•FO
RenderX

25. Fernandez RR, Huang X, Ou K. Toward a functional near-infrared spectroscopy-based monitoring of pain assessment for nonverbal patients. J Biomed Opt Oct 2017;22(10):1-12. [doi: 10.1117/1.JBO.22.10.106013] [Medline: 29076307]
26. Fernandez RR, Huang X, Ou K. A machine learning approach for the identification of a biomarker of human pain using fNIRS. Sci Rep Apr 4 2019;9(1):5645. [doi: 10.1038/s41598-019-42098-w] [Medline: 30948760]
27. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017 May 24;60(6):84-90. [doi: 10.1145/3065386]
28. He K, Zhang X, Ren S, Sun J. Harvard. 2015. Deep residual learning for image recognition URL: https://ui.adsabs.harvard.edu/ [accessed 2019-05-30] [WebCite Cache ID 78lB4pkCQ]
29. Fields H. Pain modulation: expectation, opioid analgesia and virtual pain. Prog Brain Res 2000;122:e. [Medline: 10737063]
30. Lobanov O, Zeidan F, McHaffie J, Kraft R, Coghill R. From cue to meaning: brain mechanisms supporting the construction of expectations of pain. Pain Jan 2014;155(1):129-136. [doi: 10.1016/j.pain.2013.09.014] [Medline: 24055334]
31. DaSilva A, Becerra L, Makris N, Strassman A, Gonzalez R, Geatrakis N, et al. Somatotopic activation in the human trigeminal pain pathway. J Neurosci 2002 Sep 15;22(18):8183-8192 [FREE Full text] [doi: 10.1523/JNEUROSCI.22-18-08183.2002] [Medline: 12223572]

Abbreviations

3D: 3-dimensional
AI: artificial intelligence
ANN: artificial neural network
AR: augmented reality
CNN: convolutional neural network
fMRI: functional magnetic resonance imaging
fNIRS: functional near-infrared spectroscopy
HbO: oxygenated hemoglobin
HbR: deoxygenated hemoglobin
NN: neural network
PLR: positive likelihood
S1: primary sensory

Edited by M Focsa; submitted 04.02.19; peer-reviewed by J Salisbury, S Kakarmath, Z Zhongxing; comments to author 26.03.19; revised version received 26.04.19; accepted 12.05.19; published 28.06.19.

Please cite as:
Hu XS, Nascimento TD, Bender MC, Hall T, Petty S, O'Malley S, Ellwood RP, Kaciroti N, Maslowski E, DaSilva AF
Feasibility of a Real-Time Clinical Augmented Reality and Artificial Intelligence Framework for Pain Detection and Localization From the Brain
J Med Internet Res 2019;21(6):e13594
URL: https://www.jmir.org/2019/6/e13594/
doi:10.2196/13594
PMID:31254336

©Xiao-Su Hu, Thiago D Nascimento, Mary C Bender, Theodore Hall, Sean Petty, Stephanie O'Malley, Roger P Ellwood, Niko Kaciroti, Eric Maslowski, Alexandre F DaSilva. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Comparison of Nutrigenomics Technology Interface Tools for Consumers and Health Professionals: A Sequential Explanatory Mixed Methods Investigation

Vanessa Araujo Almeida[1], RD, MSc; Paula Littlejohn[2,3], CHES, MHS; Irene Cop[3], DC, MD; Erin Brown[3,4,5], RD, MSc; Rimi Afroze[3,6,7], CMPT, BSc (Hons); Karen M Davison[3,8], RD, CHES, PhD

[1]University of Hawai'i at Manoa, College of Tropical Agriculture & Human Resources, Honolulu, HI, United States

[2]University of British Columbia, Michael Smith Laboratories, Vancouver, BC, Canada

[3]Kwantlen Polytechnic University, Department of Biology, Health Science Program, Surrey, BC, Canada

[4]Fraser Health Authority, Clinical Nutrition, Abbotsford, BC, Canada

[5]Vancouver General Hospital, Clinical Nutrition, Vancouver, BC, Canada

[6]University of Washington, School of Public Health, Seattle, WA, United States

[7]Neighborhood House Washington, Tukwila, WA, United States

[8]University of Hawai'i at Manoa, College of Social Sciences, Honolulu, HI, United States

Corresponding Author:
Karen M Davison, RD, CHES, PhD
Kwantlen Polytechnic University
Department of Biology
Health Science Program
12666 72 Avenue
Surrey, BC,
Canada
Phone: 1 604 300 0331
Email: karen.davison@kpu.ca

Abstract

Background: Nutrigenomics forms the basis of personalized nutrition by customizing an individual's dietary plan based on the integration of life stage, current health status, and genome information. Some common genes that are included in nutrition-based multigene test panels include CYP1A2 (rate of caffeine break down), MTHFR (folate usage), NOS3 (risk of elevated triglyceride levels related to omega-3 fat intake), and ACE (blood pressure response in related to sodium intake). The complexity of gene test–based personalized nutrition presents barriers to its implementation.

Objective: This study aimed to compare a self-driven approach to gene test–based nutrition education versus an integrated practitioner-facilitated method to help develop improved interface tools for personalized nutrition practice.

Methods: A sequential, explanatory mixed methods investigation of 55 healthy adults (35 to 55 years) was conducted that included (1) a 9-week randomized controlled trial where participants were randomized to receive a standard nutrition-based gene test report (control; n=19) or a practitioner-facilitated personalized nutrition intervention (intervention; n=36) and (2) an interpretative thematic analysis of focus group interview data. Outcome measures included differences in the diet quality score (Healthy Eating Index–Canadian [HEI-C]; proportion [%] of calories from total fat, saturated fat, and sugar; omega 3 fatty acid intake [grams]; sodium intake [milligrams]); as well as health-related quality of life (HRQoL) scale score.

Results: Of the 55 (55/58 enrolled, 95%) participants who completed the study, most were aged between 40 and 51 years (n=37, 67%), were female (n=41, 75%), and earned a high household income (n=32, 58%). Compared with baseline measures, group differences were found for the percentage of calories from total fat (mean difference [MD]=−5.1%; Wilks lambda (λ)=0.817, $F_{1,53}$=11.68; P=.001; eta-squared [η^2]=0.183) and saturated fat (MD=−1.7%; λ=0.816; $F_{1,53}$=11.71; P=.001; η^2=0.18) as well as HRQoL scores (MD=8.1 points; λ=0.914; $F_{1,53}$=4.92; P=.03; η^2=0.086) compared with week 9 postintervention measures. Interactions of time-by-group assignment were found for sodium intakes (λ=0.846; $F_{1,53}$=9.47; P=.003; η^2=0.15) and HEI-C scores (λ=0.660; $F_{1,53}$=27.43; P<.001; η^2=0.35). An analysis of phenotypic and genotypic information by group assignment found improved total fat (MD=−5%; λ=0.815; $F_{1,51}$=11.36; P=.001; η^2=0.19) and saturated fat (MD=−1.3%; λ=0.822; $F_{1,51}$=10.86;

XSL•FO
RenderX

P=.002; η^2=0.18) intakes. Time-by-group interactions were found for sodium (λ=0.844; $F_{3,51}$=3.09; *P*=.04; η^2=0.16); a post hoc analysis showed pre/post differences for those in the intervention group that did (preintervention mean 3611 mg, 95% CI 3039-4182; postintervention mean 2135 mg, 95% CI 1564-2705) and did not have the gene risk variant (preintervention mean 3722 mg, 95% CI 2949-4496; postintervention mean 2071 mg, 95% CI 1299-2843). Pre- and postdifferences related to the *Dietary Reference Intakes* showed increases in the proportion of intervention participants within the acceptable macronutrient distribution ranges for fat (pre/post mean difference=41.2%; *P*=.02). Analysis of textual data revealed 3 categories of feedback: (1) *translation of nutrition-related gene test information to action*; (2) *facilitation of eating behavior change*, particularly for the macronutrients and sodium; and (3) *directives for future personalized nutrition practice*.

Conclusions: Although improvements were observed in both groups, healthy adults appear to derive more health benefits from practitioner-led personalized nutrition interventions. Further work is needed to better facilitate positive changes in micronutrient intakes.

Trial Registration: ClinicalTrials.gov NCT03310814; http://clinicaltrials.gov/ct2/show/NCT03310814

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.9846

(J Med Internet Res 2019;21(6):e12580) doi:10.2196/12580

KEYWORDS

nutrigenomics; nutrigenetics; genomics; epigenomics; interface, user-computer

Introduction

Background

Since the success of the Human Genome project, scientific technology has advanced rapidly in several disciplines, including medicine and nutrition. Given that the interaction of nutrients with DNA can impact nutritional status and the development of complex diseases, nutritional genomics (nutrigenomics) has become increasingly important in nutrition practice. Nutrigenomics encompasses nutrigenetics, which investigates the effect of genetic variation on nutrient bioavailability and metabolism, and nutrigenomics, which examines how nutrients and bioactive food compounds affect human health through epigenetic modifications [1-5]. Furthermore, it forms the basis of personalized nutrition by customizing an individual's dietary plan based on the integration of life stage, current health status, and genome information. Some common genes that are included in nutrition-based multigene test panels include *CYP1A2* (rate of caffeine break down) [6,7], *MTHFR* (folate usage) [8-10], *NOS3* (risk of elevated triglyceride levels related to omega-3 fat intake) [11], and *ACE* (blood pressure response in related to sodium intake) [12].

Although the advancement of nutrigenomics-based personalized nutrition shows significant promise in improving population health, it also presents challenges. These issues include concerns about the complexity in translating gene-based results into meaningful recommendations that will lead to positive health outcomes [13-15]. Despite these issues, studies have shown that those who receive personalized nutrition interventions based on gene test results show improvements in the quality of their diet [16-19]. However, differences in dietary intakes are not always observed between *risk* and *nonrisk* groups [20], and dietary changes have not been consistently observed across all identified risk gene variants (eg, *MTHFR*) where nutrition advice is provided [16-19]. For nutrigenomics and personalized nutrition to advance in health practice, better interface educational tools (eg, Web applications and targeted messaging after personalized nutrition advice provided) need to be developed that are easily implemented by practitioners, that are understood by consumers, and that foster positive eating behavior changes. Furthermore, they need to incorporate accepted nutrition guidelines, integrate phenotypic information about current health status, and align with behavior change theory principles [5,20]. This study proposes to compare standard and tailored personalized nutrition approaches based on gene testing and to elicit participant feedback about their experiences with the 2 types of interventions. The study results are intended to generate data that will improve nutrigenomics-based education tools for consumers and health professionals.

Objectives

The main study objectives were to compare a practitioner-facilitated personalized dietary approach that uses genotypic and phenotypic information with a self-driven approach and their impacts on changing participant's knowledge, motivation, and behavior related to eating habits, the quality of their diet, and the quality of their life. It was hypothesized that significantly higher levels of knowledge, motivation, behavior, and quality of life would be reported and that there would be a higher level of diet quality changes in the group that receives personal DNA diet information plus customized dietary advice (practitioner led) compared with the group that is provided personal DNA diet information (direct-to-consumer, self-driven approach) only. In addition, self-efficacy was evaluated as a potential mediator/moderator of dietary changes and outcomes. Focus group interview data were also collected to facilitate interpretation of the intervention data.

Methods

Study Design

The sequential explanatory mixed methods investigation consisted of (1) a randomized controlled trial (2:1 allocation ratio) comparing standard self-driven versus practitioner-facilitated approaches that use DNA-based diet information and (2) qualitative investigation of participants'

XSL•FO
RenderX

experiences to help interpret the intervention's quantitative outcomes. The study protocol, including paper-based or Web-based data collection forms, was approved by Quorum Institutional Review Board (protocol #32220CDN/1). All the participants provided initial Web-based consent to collect eligibility screening information and, if eligible, baseline information related to their health. At the participant's first site visit, a second written consent form was reviewed, which participants signed to confirm their continued involvement in the study. The protocol was registered with the US National Library of Medicine (trial registration #NCT03310814) and is also detailed elsewhere [21].

Participants and Setting

The participants selected for the study included healthy medically stable adults (aged 35 to 55 years) residing in the greater Vancouver area of the province of British Columbia, Canada. The participants were recruited via social media, a newspaper article, and posters. The eligibility criteria included that they wanted to improve their health, could understand and provide informed consent, and were willing to provide a buccal swab for DNA testing. The exclusion criteria are specified in Multimedia Appendix 1.

Description of Study Groups

The participants were randomized using a random number generator by a statistician independent to the study into either the intervention or control (C) group. The intervention (I) group received their gene test result report (standard) plus a personalized nutrition plan that integrated information about their gene test results, health information, personal goals, and dietary intakes (based on nutrient analysis of 3-day food records collected at baseline). Their DNA report results and personalized nutrition plan were reviewed with a trained research registered dietitian (RD) who counseled them about the DNA-based diet recommendations derived from information about their gene markers and variants. For example, if a person possessed a dietary fat utilization–based genotype that was associated with increased risk of a health outcome (eg, dyslipidemia), they were provided *targeted* directives about types (eg, omega-3 fatty acids) and amounts of fat intake (eg, try to eat less than 80 g of fat per day). If a participant did not possess the specified dietary fat utilization–based risk variant, they received the standard *Dietary Reference Intakes* –based recommendations [22,23]. Where possible, recommendations based on the gene test results were clustered to provide manageable and meaningful dietary guidance to the participant (Table 1). As the final step in nutrition consultation, the RD worked collaboratively with the participant to define up to 3 targeted nutrition-related goals that were entered into the Web-based nutrition assessment questionnaire and written in the personalized nutrition report. Both groups received 3 follow-up emails (1 every 2 weeks post intervention) with information about gene-based personalized nutrition as well as tips and reminders (eg, information about food label reading) to help them reach their nutritional goals. After the study's completion, participants in the C group had the option to have dietary counseling from the research RD.

Table 1. Description of gene test.

Component measured	Nutrient or food subcomponent tested[a]
Diet management	Carbohydrates[b]; cholesterol—high density lipoprotein and low density lipoprotein[c]; fat—dietary, stored, monounsaturated, and saturated[b]; insulin sensitivity; and protein[d]
Weight response	Body mass index
Food tolerances	Alcohol; caffeine; gluten; lactose; salt[e]; and sugar craving[f]
Food taste and preferences	Caffeine preference; carbohydrate[b] preference; fat preference[c]; protein preference[d]; and taste: bitter, salt[e], and sweet[f]
Vitamins, minerals, and essential fats	Vitamins A, B_6, B_9 (folate), B_{12}, C, D[g], and E; calcium[g]; iodine; iron; and omega 3 and 6[c]

[a]Dietary advice was clustered by subcomponents labeled with the superscripted letters b to g.

[b]Carbohydrate-related advice.

[c]Fat-related advice.

[d]Protein-related advice.

[e]Salt-related advice.

[f]Sugar-related advice.

[g]Vitamin D and calcium-related advice.

Study Visits

The 4-month study consisted of 6 stages.

Recruiting/Screening

This included the initial eligibility screen and, where applicable, baseline assessment (Web based) that collected information about sociodemographics; current health status (eg, presence of health conditions and medication and supplement usage); quality of life; self-efficacy; questions about knowledge, motivation, and action related to DNA-based information; stage of change; physical and sedentary activities; food intakes (food frequency and food selection); anthropometrics; and sleep quality. In addition, participants were sent 3-day food records to complete within 7 (±3 days) of their first site visit.

XSL•FO
RenderX

Baseline Assessment

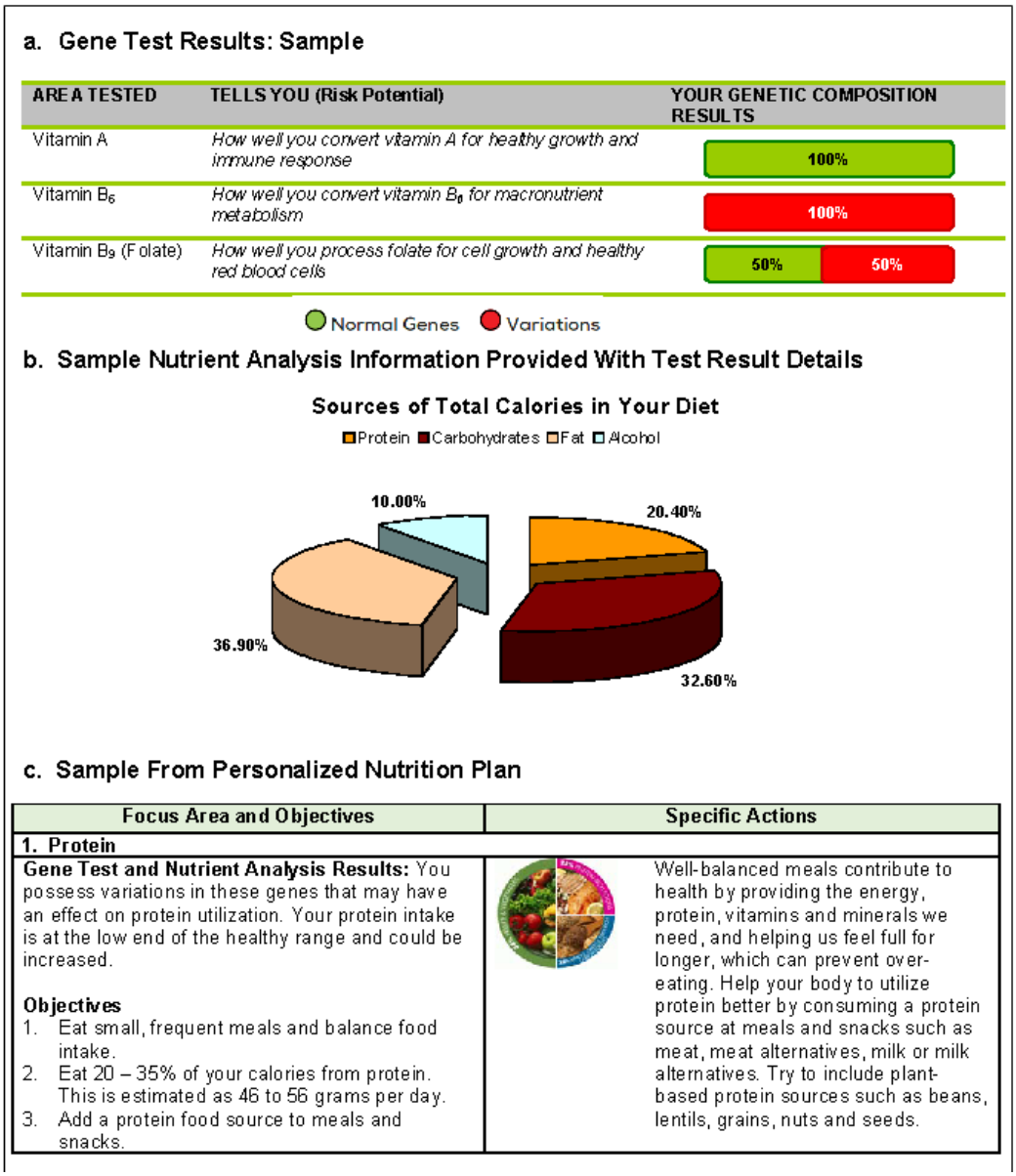
A site visit was conducted with the RD who reviewed the participant's baseline health assessment information and food records; collected a buccal cheek swab sample; and measured the participant's height, weight, and waist and hip circumference according to standardized protocols [24].

Bar-coded buccal DNA samples were collected using Oracollect-DNA OCR-100 swabs (DNA Genotek) and stored between 15° and 30°C. Participants did not eat or drink at least 30 minutes before buccal sample collection. The samples were analyzed using Agena MassArray at the Clinical Genomics Centre at Mount Sinai Hospital. The gene tests included 5 evidence-based components (Table 1) [25].

Personalized Nutrition Visit or Standard Report Distributed

I group participants came to the site to have their personalized nutrition report reviewed (see Figure 1 for samples from the interface tools); those in the C group received an email with the standard report.

Figure 1. Examples from educational tools.



Follow-Up #1

A Web-based survey was sent to I and C participants at week 3 post intervention to collect data about any changes in income; social support; and knowledge, behavior, and action; stage of change; and adverse event information. For the I group, questions that asked about whether knowing one's personal DNA helped with eating behavior change were also included.

Follow-Up #2

At 6 weeks after the intervention, the participants received the Web-based baseline health assessment questionnaire and food records to complete in preparation for the final onsite visit (week 8 post intervention). Additional questions were asked about whether having the diet-based DNA information led to action in areas such as label reading and making healthier food choices when eating out, at the grocery store, and at home.

XSL•FO
RenderX

Final Visit

The final site visit (I and C groups) included a review of the participant's health assessment and food record information. Anthropometric measures were repeated. The C-group participants received their individualized gene test–based diet plans and had the option to book appointments with the RD to have them reviewed. Focus groups were then conducted to solicit feedback about participants' responses to their gene test results and RD consultation as well as about barriers, facilitators, and targets for improvement related to nutrigenomics-related education.

Data Collection Tools and Outcome Measurements

FluidSurveys software [26] was used to construct the Web-based closed questionnaires. The Web-based questionnaires were developed using the Checklist for Reporting Results of Internet E-Surveys [27], standard measurements (detailed in measurements section), and protocols for nutrition assessment [25]. All questionnaires contained 12 or less screens (pages) and were pilot tested with study staff and student volunteers (n=11) to assess the usability and technical functionality. The participants were emailed instructions and the links to each Web-based questionnaire at the appropriate times during the study. Quality control data collection procedures included ensuring all questions in each questionnaire were completed, following up with participants about responses where required, and deleting any duplicate entries according to internet protocol addresses. All the data were collected and stored in accordance with Quorum institutional review board guidelines. Copies of the questionnaires may be obtained by emailing the corresponding author. The outcome measures are described in the following sections.

Outcomes

Nutrition-Related Outcomes

All dietary intake data collection was done according to standard procedures [24,28]. The 3-day food records measured pre/post caloric, macronutrient, micronutrient, and food group intakes, which were compared with nutrition standards (eg, *Eating Well with Canada's Food Guide* and *Dietary Reference Intakes*). The Canadian version of the Healthy Eating Index (HEI-C) [29] was used to assess diet quality. The Web-based questionnaires also included food selection questions about the types of food consumed, dietary restraint, food insecurity, motivation to change diet, and eating behavior changes. These were based on validated measures [30,31] and current literature related to measuring motivation and eating behavior changes.

Health-Related Quality of Life Short Form 8

The Health-Related Quality of Life Short Form 8 is a validated measurement tool of quality of life, functional health, and well-being [32] based on a 4-week recall period. Each item has a 5- or 6-point response range. Higher summary physical and mental component scores indicate better health [33].

General Self-Efficacy

General Self-Efficacy (GSE) is a validated item measure of self-efficacy shown to correlate with emotion, optimism, and work satisfaction [34]. It contains 10 questions each with 4 categories of responses that include not at all true, hardly true, moderately true, and exactly true. The scores for each question range from 1 to 4 and the total score is calculated by finding the sum of all items. The scale score ranges between 10 and 40; higher scores indicate greater self-efficacy. The GSE has high reliability, stability, and construct validity that have been confirmed (Cronbach alpha ranges from .86 to .94) [35].

Measures of Change in Knowledge, Motivation, and Behavior

Overall, 3 questions to assess for changes in knowledge, motivation, and behavior related to DNA-based dietary advice were developed by the authors based on the Stages of Change Model [36] and current review of the evidence. They were pilot tested among 11 young adults (aged 20 to 40 years) to assess for comprehension and face validity.

Covariates

The covariates related to dietary intake and health behavior that were measured included the following.

Natural Health Products Usage

Data about natural health products (NHP; eg, micronutrients and botanicals) included type (natural product number), dose, and frequency of use. Participants who were taking NHPs were advised at baseline to keep the type, dose, and frequency the same throughout the study.

Physical and Sedentary Activities

The physical activity index [37] collected data about the frequency, duration, and intensity of participation in certain activities in the previous 3 months to derive a score that represented the average daily energy expended on leisure time physical activity. A composite measure of sedentary activities was based on total screen time in the previous 7 days [38].

Sleep Quality

The Patient-Reported Outcomes Measurement Information System Sleep Disturbance scale-short form [39] was used to assess sleep quality over the previous 7 days. Individual scale items are scored from 1 to 5 and the total scores range between 8 and 40. Lower scores represent a lesser degree of sleep-related impairments.

Stress

Measures of stress were based on 2 validated questions from the Canadian Community Health Survey [40].

Anthropometrics

Body mass index (kg/m^2) and waist-to-hip ratio were calculated from standardized height, weight, waist circumference, and hip circumference measures [24].

Sociodemographics

Measures of sex/gender, age, relationship status, income, race/ethnicity, and perceived social support were included.

XSL•FO
RenderX

Adverse Events

At each study contact point, the participants were asked if they experienced any adverse events or change in and/or had started any new medications or NHPs.

Data Analysis

Quantitative Analysis

Food Intake and Nutrient Analysis

Nutrient analysis was conducted using ESHA—The Food Processor Nutrition Analysis and Fitness software and the Canadian Nutrient File [41,42]. Averages of the 3 days of nutrient values were used in the analysis. The food frequency questionnaire (FFQ) values were used to derive usual intakes of the nutrients of interest [24].

Descriptive and Inferential Analysis

Means (SDs) or medians (and interquartile range) were reported based on a given continuous variable's distribution. Subject characteristics, group comparisons, and pre- and postintervention differences were analyzed using the Student *t* tests, binomial tests of 2 proportions, Fisher exact tests, and 2-way repeated-measures analysis of variance with Bonferroni post hoc tests, where appropriate. All analyses were done on an intent-to-treat basis using STATA software [43].

Qualitative Analysis

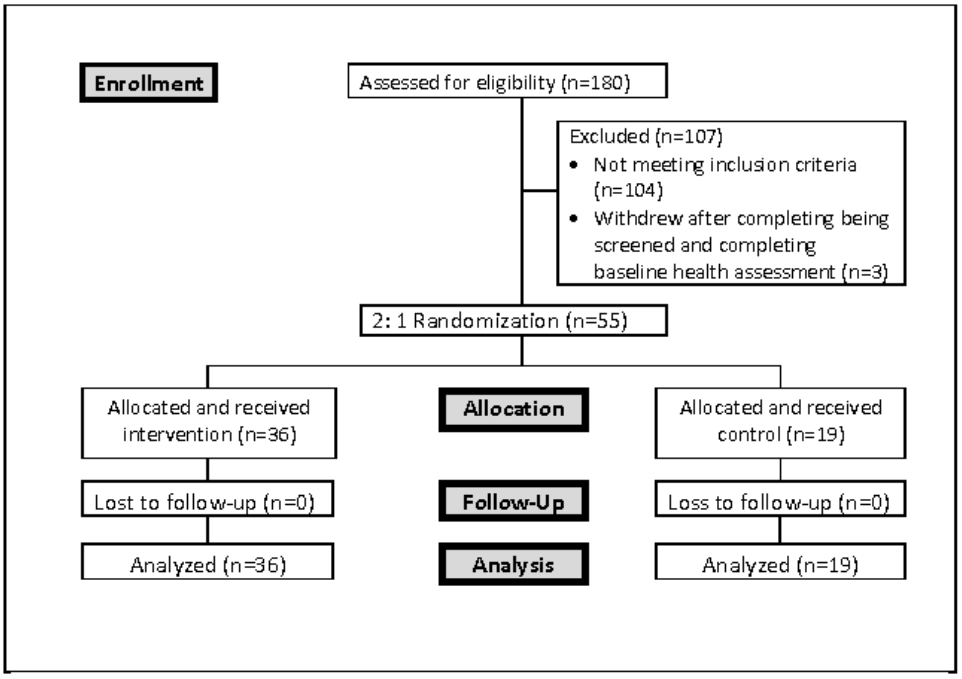
Textual data from the Web-based questionnaires (eg, participant's personal dietary goals) were grouped into categories where feasible. Data from the focus groups were transcribed, entered into NVivo [44], and analyzed by research team members using interpretative thematic analysis to identify patterns, concepts, themes, and examples in relation to existing behavior change theories and the study objectives [45]. Interpretations were reviewed by research team members and participants to check for descriptive and interpretative validity. Qualitative data were reported based on thematic analysis derived from 3 independent reviews of the textual data.

Results

Sample

A total of 478 persons expressed interest in study participation. These individuals were invited sequentially from the recruitment list and balanced by sex/gender, where possible. However, the ratio of females to males who were interested in the study was approximately 8:1, and although all males on the recruitment list were contacted, there was an imbalance in the male/female participant ratio. Of the 478 individuals, a total of 180 (38%) were invited to complete the Web-based eligibility screening questionnaire (Figure 2); 73 of the invited individuals (73/180, 41%) were deemed eligible. Of the 73 eligible individuals, 58 enrolled and 55 (95%) completed the baseline health assessment questionnaire and food records (55/73, 75%). The final sample consisted of 55 adults aged between 37 and 57 years (mean 45.8 years, SD 5.8). Most participants were women (n=41, 74.6%), had graduated from a university or college (n=34, 62%), earned an income higher than Can $90,000 (n=32, 58.8%), and were in a relationship (n=47, 86%). Almost 51% (n=28) of the participants believed they were eating a healthy diet, about two-thirds indicated their health status was *good* (n=25, 46%) or *very good* (n=14, 26%), and 91% (n=50) reported they had favorable social support networks. There were no significant differences between the I and C groups based on various baseline characteristics (Table 2). In addition, the proportion of participants with the gene risk variants did not differ between the I and C groups (proportion range 6% to 69%; *P* values range from .27 to .77).

Figure 2. Participant selection.



XSL•FO
RenderX

Table 2. Description of sample—baseline characteristics.

Demographics and health-related factors	Intervention (n=36)	Control (n=19)	*P* value
Age (years), mean (SD)	45.4 (6.3)	46.6 (5.0)	.49
In a relationship, n (%)	29 (81)	18 (95)	.47
Postsecondary education, n (%)	20 (56)	14 (74)	.35
Income above Can $90,000, n (%)	19 (53)	13 (68)	.13
Inactive, n (%)	20 (56)	14 (74)	.35
General Self-Efficacy, mean (SD)	32.8 (4.0)	32.7 (4.4)	.88
Health-related quality of life, mean (SD)	62.6 (12.0)	64.6 (9.5)	.53
Body mass index, mean (SD)	28.7 (5.5)	28.1 (4.2)	.69

Changes in Knowledge, Attitudes, and Behavior

At baseline, more than 90% of respondents (minimum 51) indicated that the diet-related gene test information helped them to understand their health better as well as motivated them to take action related to healthy eating (Multimedia Appendix 2). The proportions of individuals who agreed to statements about taking action on reading nutrition labels more often, selecting healthier food choices at restaurants, purchasing healthier food, making healthier meals at home, taking supplements to support their DNA, and adjusting their eating according to their DNA information were high across both groups; consistently higher proportions were found in the I group. Responses to Stages of Change questions indicated that those in the I group were more likely to report the intention of making healthy changes to their diets within the next month (Multimedia Appendix 2). Responses to questions given to the I group about diet-related changes that were asked at 3 weeks postintervention indicated that 31% (n=11) starting to take supplements, 81% (n=29) took diet-specific actions, and 42% (n=15) indicated they adjusted their eating according to their DNA information.

Changes in Dietary Intake, Anthropometrics, Self-Efficacy and Quality of Life

Significant pre-/postintervention differences (Table 3) were found for the percentage of calories from total fat (mean difference [MD]=−5.1%; Wilks' lambda [λ]=0.817; $F_{1,53=}$ 11.68; *P*=.001; eta-squared [η^2]=0.183) as well as from saturated fat (MD=−1.7%; λ=0.816; $F_{1,53}$=11.71; *P*=.001; η^2=0.18) and health-related quality of life (HRQoL) scores (MD=8.1 points; λ=0.914; $F_{1,53}$=4.92; *P*=.03; η^2=0.086). There were significant differences between groups over time for sodium (λ=0.846; $F_{1,53}$=9.47; *P*=.003; η^2=0.15) and HEI-C scores (λ=0.660; $F_{1,53}$=27.43; *P*<.001; η^2=0.35; Figure 3).

When group assignment was stratified by phenotypic plus genotypic information, improved total fat (MD=−5%; λ=0.815; $F_{1,51}$=11.36; *P*=.001; η^2=0.19) and saturated fat (MD=−1.3%; λ=0.822; $F_{1,51}$=10.86; *P*=.002; η^2=0.18) intakes were indicated. In addition, significant differences between groups over time for sodium (λ=0.844; $F_{3,51}$=3.09; *P*=.04; η^2=0.16) with post hoc analysis showing pre/post differences for those in the I group that did have the risk variant (premean 3611 g, 95% CI 3039-4182; postmean 2135 g, 95% CI 1564-2705) and did not have the risk variant (premean 3722 g, 95% CI 2949-4496; postmean 2071 g, 95% CI 1299-2843). Improvements in omega 3 fatty acid intakes were close to significant (MD=0.34 g; λ=0.926; $F_{1,51}$=3.99; *P*=.051; η^2=0.07). Pre/post differences related to *Dietary Reference Intakes* showed increases in the proportion of intervention participants within the acceptable macronutrient distribution ranges for fat (pre/post difference=41.2%; *P*=.02).

For self-efficacy and quality-of-life measures, there were significant group pre/post intervention differences for HRQoL scores only (MD=8.1 points; λ=0.914; $F_{1,53}$=4.92; *P*=.03; η^2=0.086).

XSL•FO
RenderX

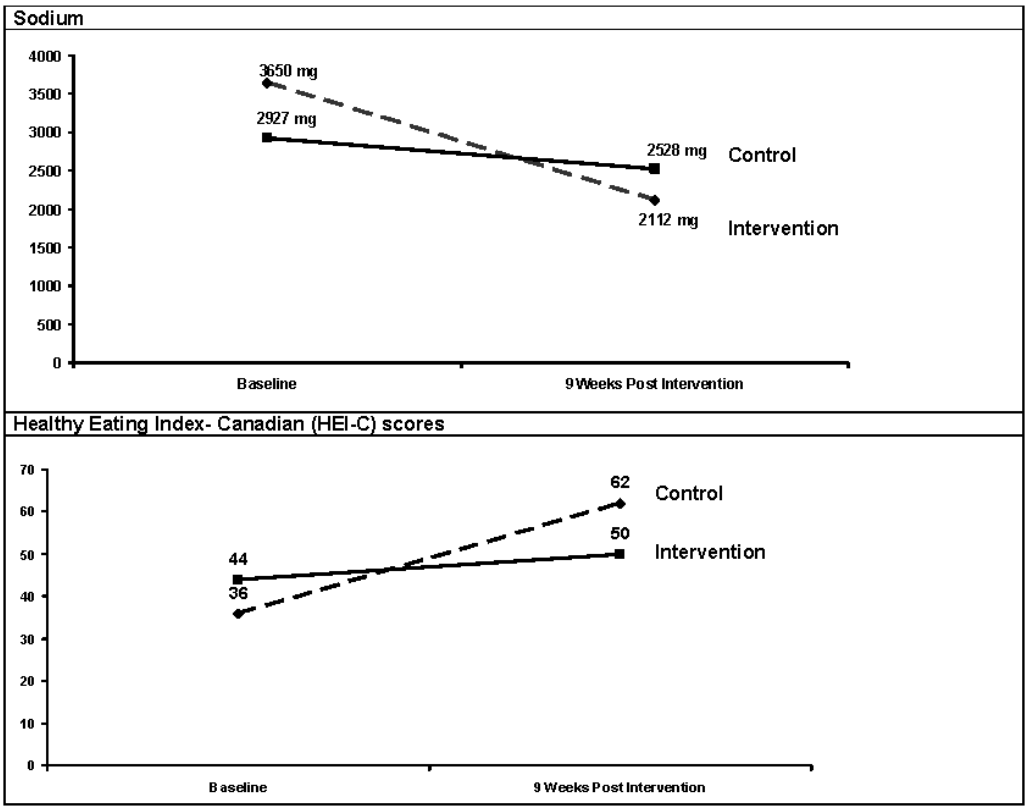
Table 3. Pre- and postintervention comparisons of nutrient intakes, overall diet quality, Health-related Quality of Life Scores, and General Self-Efficacy scores.

Nutrient (units)	Mean difference		Wilks' lambda (λ)	*F* (df)[a]	eta-squared (η^2)	*P* value
	Intervention	Control				
Macronutrients and food components						
Fat (% total calories)	−5.1	−2.9	0.817	11.68 (1,53)	0.183	.001
Saturated fat (% total calories)	−1.3	1.6	0.816	11.71 (1,53)	0.184	.001
Omega 3 (g[b])	−0.4	0.4	1.000	0.02 (1,53)	0.000	.88
Omega 3 (% total calories)	0.1	0.3	0.964	1.93 (1,53)	0.036	.17
Omega 6 (g)	−2.2	−3.4	0.688	23.58 (1,53)	0.312	<.001
Omega 6 (% total calories)	0.8	0.6	0.908	5.26 (1,53)	0.092	.03
Sodium (mg[c])	−1537.9	−398.6	0.846	9.47 (1,53)	0.150	.003[d]
Sugar (g)	1.8	−0.1	0.979	1.11 (1,53)	0.021	.30
Fiber (g)	−1.2	−1.2	0.000	9.47 (1,53)	0.000	.99
Vitamins and minerals						
Vitamin B_6 (mg)	−1.6	0.4	0.989	0.58 (1,53)	0.011	.45
Vitamin B_9 (mcg DFE[e])	−89.1	−57.0	0.950	2.72 (1,53)	0.050	.11
Vitamin B_{12} (mcg[f])	-0.4	−1.7	0.883	6.90 (1,53)	0.117	.011
Vitamin D (IU[g])	1.2	2.2	1.000	0.01 (1,53)	0.000	.92
Calcium (mg)	−176.7	−121.9	0.899	5.81 (1,53)	0.101	.02
Diet quality						
Healthy Eating Index- Canadian	26.5	6.9	0.655	27.34 (1,53)	0.345	<.001[d]
Other outcomes						
Health-related quality of life short form-8	7.0	5.3	0.914	4.92 (1,53)	0.086	.03
General Self-Efficacy	0.9	0.1	0.975	0.01 (1,53)	0.001	.27

[a]Exact statistic.
[b]g: gram.
[c]mg: milligram.
[d]Significant interactions of time by groups.
[e]DFE: dietary folate equivalent.
[f]mcg: microgram.
[g]IU: international unit.

XSL•FO
RenderX

Figure 3. Time-by-group interactions for sodium and diet quality scores.



Qualitative Analysis

On the basis of analysis of textual data from the Web-based questionnaires, the personal goals that participants set for themselves tended to be related to weight loss; reducing sugar, fat, and processed food intakes; and increasing consumption of fruits and vegetables. In total, 5 focus groups (with a range of 5 to 9 participants) were conducted that generated more than 36,000 words of textual data. From the analysis, 3 categories of feedback emerged: (1) interpretation of nutrition-related gene test information, (2) facilitation of eating behavior change, and 3) directives for personalized nutrition practice.

Category 1: Interpretation of Nutrition-Related Gene Test Information

Participants from both groups thought richly detailed reports containing their nutrition-related gene test results were valuable:

> *...I think this is hugely valuable I really like this and I think you know inflammation seems to be a very big problem for a lot of people that could definitely help you find the right proteins the right fits whatever it is that it's going to assist your body in lowering that level and the next one hopefully having less disease or less arthritis or issues.*
>
> *I think it was the detail. I was like "Wow" because I am scrolling down and I am like there is more and more. And you know, kind of did a cursory look and then I started to look more and I was like "Wow" because you know it really gets into specifics.*
>
> *I loved it. It's well designed and well-structured so it's pretty good. Color coding was really great and you can easily get it just by a glance and see exactly what you are looking at.*

Both groups shared insights about challenges with interpreting their results. Some C-group participants (n=4; received report only) tried to find a practitioner to review the results, which lead to variable responses:

> *I tried to hire a dietitian or nutritionist and I told them I had the DNA results and I said, "Are you willing to look at them and making a meal plan?" And they both said no.*
>
> *Umm well my first impression was that some of the recommendations seemed to be non specific to but I just filtered through those and focused in on the ones that I thought were those were clearly for me...*
>
> *...there are ideas in there but I needed it to be spelled out and I was willing to pay someone but I just couldn't find anyone to do it who was willing. That's the only thing.*
>
> *...they {results} also gave me an opportunity to sit down with my Dr and talked about what I've been feeling and what these results and let's play with this and so she thought this was awesome. She thought this was really good she said...She's very innovative and forward thinking.*

Category 2: Facilitation of Eating Behavior Change

The participants revealed different components of their diet they either felt motivated to act on or that they took action on, regardless of which group they were assigned to. They expressed

XSL•FO
RenderX

that the tailoring of information specific to their individual circumstances was motivating:

> *I think it made me feel more empowered. You just have more information.*
>
> *In the past, I would be making all my selections and choices based on the general knowledge of a good lifestyle but now I know more about me and so I make personal choices. Knowing more about yourself and your body and what choices you make would be beneficial for you and what would be the most useful.*
>
> *...I increased my protein and tried to make sure to incorporate some protein every meal and instead of just going for a quick grab of carb stuff I'd made sure that there was a balance, there was a protein...*
>
> *...orange vegetables too so I'm more conscious about going after those as well and processing of the proteins I'm consuming too...*

For those in the I group, the addition of the personalized nutrition report and dietitian consult that included review of nutrient analysis information from food records was reported to be helpful in selecting dietary intake targets to change:

> *It's amazing that the information that [dietitian's name] gave me just learned about different oils and how they interact...*
>
> *...I think that having [dietitian's name] discuss it hit home more and it makes me more conscious, more aware. Yeah it also helps your understanding level because sometimes you look at things and your brain doesn't connect with it the same way as having [dietitian's name] say "well if you did this or that" and then your brain goes "oh okay, I can get that" and then you commit to it better.*
>
> *I think it was fantastic especially if you are doing a collaboration with someone you could physically talk to and it would be more meaningful...having somebody to actually talk to and work through it together and say, "Okay, what does this all mean?" or "How does this all look?" really validates and reinforces the information that is contained in the book.*
>
> *I think it did some motivation for sure, but I am not going to say 100% because it's not like I changed my whole [lifestyle]. You know, I choose salad over my French fries most of the week...I wouldn't say it was the genetic report that made me change this all, but it was the food record that made me realize it was out of whack.*
>
> *...I think the benefits personally to me it was also being able to sit down and talk to [dietitian's name] yes in relation to the report...*
>
> *I didn't realize how many hidden fats I was eating with the curries and all that but writing it down on paper and then have to have a talk about it. And the recommendation try to replace the fats with proteins so I am more likely to reach for the nuts than anything else...*

Category 3: Directives for Future Personalized Nutrition Practice

Consistent with previous literature [13-15], many participants stated factors such as cost and privacy concerns were potential barriers to having nutrition-related gene testing done. Others within the I group noted that ongoing reinforcement was particularly helpful to them when trying to take action on the results. Most participants in both the I and C groups discussed that the major nutrients and salt were their key targets as the advice was easy for them to remember and follow; food components that were frequently mentioned in the focus groups by both the I and C groups included fat, caffeine, and lactose. Many indicated that trying to translate gene-based micronutrient results and the associated food-based recommendations into their daily diet was much more difficult:

> *But it didn't kick in again when I got the reminder on the survey and because I felt lost when we were left on our own whether it was sticking to it or not...And then we got the reminder, oh yeah yeah yes I did oranges...So they kicked me back into action...*
>
> *I knew the vitamins one was probably the least influential on me and of least interest to me because I couldn't wrap my head around how much all of these different types of vitamins that I am consuming and whether or not if I truly understand what each one what it's purpose is and how it serves me. So I probably spent the least amount of time trying to, you know, contemplate and reflect on those results.*

Discussion

Principal Findings

This study compared the effectiveness of standard and tailored gene-based personalized nutrition interventions in improving knowledge and motivation to change eating behavior, dietary intakes, and quality of life. A second objective was to elicit participant feedback from focus groups to better understand the quantitative findings. The results indicated that over a 9-week period, tailored interventions contributed to improved fat and sodium intakes, overall diet quality, and quality of life. In addition, the participants receiving these types of interventions typically indicated that they intended to make changes to their diet in the near future. When phenotypic plus genotypic information by group assignment was considered, improved total fat, saturated fat, and sodium intakes were found among those who possessed the risk genotype and received tailored dietary advice. Qualitative findings revealed that participants thought that the gene-based nutrition information motivated them to change eating behaviors. However, many indicated they found some of the information complicated, particularly the translation of micronutrient-based results and associated food-based recommended actions.

Results Compared With Existing Research Literature

This study's findings were consistent with other investigations that showed greater improvements in diet quality after receiving personalized practitioner-facilitated nutrition interventions when compared with receiving a standard gene test report [16-19,44].

XSL•FO
RenderX

Our findings indicated that at 6 weeks postintervention, many participants indicated that the DNA information related to diet continued to motivate them and to take action on healthy eating. However, the proportion of individuals reporting this declined over the course of the study indicating that ongoing practitioner-led reinforcement may be more helpful in keeping participants focused on continued action to eat healthier. Similar to other studies, when genotypic and phenotypic information was considered by group assignment, significant positive differences in select nutrient intakes occurred [44]. These findings that demonstrate that the provision of nutrition-based genetic information improves dietary behaviors are practices that can facilitate eating behavior changes [45]; however, further research is needed to determine their generalizability to other populations and if the improvements are sustained long term.

The results that suggested receiving gene test results motivated participants to change their diet have also been reported in other studies [46]. In addition, the qualitative findings suggested that for many participants their main diet-related goal after receiving the results was weight management. Research suggests that motivation to positively change eating behaviors is also based on factors such as ethical concerns and mood. If a practitioner has previous knowledge of what people's motives are, they can use this information combined with gene test results to develop appropriate goals and further inspire them to make positive dietary changes [19].

Based on the experience of other investigators [16-19], the direction in dietary changes in relation to the nutrition-based *risk* variant education provided was not always consistent, and in some instances, they were counterintuitive. An example of this included the significant declines in MDs of intake for calcium and vitamin B_{12}. It is thought these decreases may have been because of reductions in intake of sources of saturated fat (eg, meat and milk-based products). Alternatively, as found in the Food4Me Study [18], if people were informed that they had the nonrisk version of the genotype, they may not be motivated to focus on the associated dietary component. The findings that showed declines in folate intake MDs are consistent with other findings, which suggest that regardless of whether one is told they do or do not have the risk version of the *MTHFR 677CT* gene, the information does not seem to translate into meaningful differences in folate intake measures [47]. On the basis of insights from our qualitative data, the variable findings among intakes of vitamin and mineral intakes may be because of a lack of understanding by the participant of their functions and relevance. Clearly, future work is needed to better understand how the presentation of information for different genotypes can be better used to improve a targeted eating behavior.

To the best of our knowledge, the findings that showed time-by-group interactions for sodium and overall diet quality have not been reported elsewhere. In particular, the I group had more pronounced reductions in dietary sodium intake compared with the C group. These results have implications for the prevention and treatment of hypertension, a condition that can be exacerbated by high sodium intakes and can progress to coronary heart disease in some individuals. Hypertension is becoming increasingly prevalent, affecting an estimated 1.13 billion people worldwide [48,49]. Hypertension and coronary heart disease are related to both genetic predisposition and environmental exposures such as an unhealthy diet. This study's findings imply that tailored nutritional advisement could help prevent epigenetic changes that contribute to the development and progression of these chronic conditions. As a result, substantial savings in health care costs could be realized.

The findings that indicated that significantly greater improvements in quality-of-life measures occurred in the I group were novel. Although previous studies have shown relationships between poorer quality diet and worse quality of life [50], to the best of our knowledge, this is the first time this has been reported related to gene-based personalized nutrition interventions. We thought that GSE might play a role in this relationship; however, when we separately analyzed dietary quality as a dependent variable and quality-of-life scores as the independent variable and controlled for GSE, the findings were nonsignificant.

It was surprising that dietary improvements were not observed across all results where nutrition-related risk variants were present in the participants. The focus group data helped to provide insights about these findings by suggesting that although participants found the information useful, it was also overwhelming. As a result, they would try to simplify the implementation of their results by adjusting intakes of the major nutrients and sodium and pay less attention to the micronutrients-based gene information. However, in doing this, they may not be substantially improving the overall quality of their diet or meeting all of their micronutrient requirements. For the future, improved nutrition education strategies are needed that will facilitate the uptake of all recommended dietary changes that are based on gene test results.

Strengths and Limitations

This study's strengths include its high retention rate, focus on a defined adult population, assessment of dietary intakes using FFQ estimates and 3-day food records to reduce misreporting error, and the provision of quantitative and qualitative data. However, this investigation could have been strengthened by including objective measures such as biochemical indicators of nutrient status. The modest sample size prevented stratification of results based on individual genes. Furthermore, given the composition of the sample was mainly female and Caucasian, the generalizability of the results is limited.

Implications and Conclusions

Providing participants with DNA information related to diet improved knowledge, motivation, and action related to healthy eating. However, tailored practitioner-led, gene-based personalized nutrition interventions tend to be more effective in improving dietary intakes of key target nutrients such as fat and sodium. Further work is needed to produce more comprehensive changes in dietary intakes that are based on gene test results. The study results will be leveraged to generate new, tailored, and digitally based nutrigenomics education tools that may help to advance gene-based personalized nutrition practice.

XSL•FO
RenderX

Acknowledgments

This study is funded by a Natural Sciences and Engineering Research Council of Canada Engage Grant (Grant #500692 – 16). The authors would like to acknowledge assistance from Charvi Gulati, Sydney Gloanec, and Breshna Saidiy during the focus group interviews and their assistance with nutrient analysis. KMD and VAA were partially supported by the Fulbright Program. RA was support by the Work Study Program.

Authors' Contributions

VAA drafted the manuscript and all other authors reviewed and provided feedback with each draft.

Conflicts of Interest

PL provides Web-based nutrigenomics education. No other conflicts of interest are declared.

Multimedia Appendix 1

Participant exclusion criteria.

[DOC File, 30KB - jmir_v21i6e12580_app1.doc]

Multimedia Appendix 2

Knowledge, attitudes, and behavior baseline to study completion. Stages of change: proportion who changed throughout the study.

[DOC File, 100KB - jmir_v21i6e12580_app2.doc]

Multimedia Appendix 3

CONSORT‐EHEALTH checklist (V 1.6.2).

[PDF File (Adobe PDF File), 110KB - jmir_v21i6e12580_app3.pdf]

References

1. Simopoulos AP. Nutrigenetics/Nutrigenomics. Annu Rev Public Health 2010;31:53-68. [doi: 10.1146/annurev.publhealth.031809.130844] [Medline: 20070200]
2. De Caterina R. Genetic and nutritional interactions in cardiovascular disease. World Rev Nutr Diet 2011;102:150-155. [doi: 10.1159/000327803] [Medline: 21865829]
3. Tammen SA, Friso S, Choi S. Epigenetics: the link between nature and nurture. Mol Aspects Med 2013 Jul;34(4):753-764 [FREE Full text] [doi: 10.1016/j.mam.2012.07.018] [Medline: 22906839]
4. Berná G, Oliveras-López MJ, Jurado-Ruíz E, Tejedo J, Bedoya F, Soria B, et al. Nutrigenetics and nutrigenomics insights into diabetes etiopathogenesis. Nutrients 2014 Nov 21;6(11):5338-5369 [FREE Full text] [doi: 10.3390/nu6115338] [Medline: 25421534]
5. Mathers JC. Nutrigenomics in the modern era. Proc Nutr Soc 2017 Dec;76(3):265-275. [doi: 10.1017/S002966511600080X] [Medline: 27819203]
6. El-Sohemy A, Cornelis MC, Kabagambe EK, Campos H. Coffee, CYP1A2 genotype and risk of myocardial infarction. Genes Nutr 2007 Oct;2(1):155-156 [FREE Full text] [doi: 10.1007/s12263-007-0043-4] [Medline: 18850169]
7. Palatini P, Ceolotto G, Ragazzo F, Dorigatti F, Saladini F, Papparella I, et al. CYP1A2 genotype modifies the association between coffee intake and the risk of hypertension. J Hypertens 2009 Aug;27(8):1594-1601. [doi: 10.1097/HJH.0b013e32832ba850] [Medline: 19451835]
8. Solis C, Veenema K, Ivanov AA, Tran S, Li R, Wang W, et al. Folate intake at RDA levels is inadequate for Mexican American men with the methylenetetrahydrofolate reductase 677TT genotype. J Nutr 2008 Jan;138(1):67-72 [FREE Full text] [doi: 10.1093/jn/138.1.67] [Medline: 18156406]
9. Guinotte CL, Burns MG, Axume JA, Hata H, Urrutia TF, Alamilla A, et al. Methylenetetrahydrofolate reductase 677C-->T variant modulates folate status response to controlled folate intakes in young women. J Nutr 2003 May;133(5):1272-1280. [doi: 10.1093/jn/133.5.1272] [Medline: 12730409]
10. Tsang BL, Devine OJ, Cordero AM, Marchetta CM, Mulinare J, Mersereau P, et al. Assessing the association between the methylenetetrahydrofolate reductase (MTHFR) 677C>T polymorphism and blood folate concentrations: a systematic review and meta-analysis of trials and observational studies. Am J Clin Nutr 2015 Jun;101(6):1286-1294. [doi: 10.3945/ajcn.114.099994] [Medline: 25788000]

XSL•FO
RenderX

11. Ferguson JF, Phillips CM, McMonagle J, Pérez-Martínez P, Shaw DI, Lovegrove JA, et al. NOS3 gene polymorphisms are associated with risk markers of cardiovascular disease, and interact with omega-3 polyunsaturated fatty acids. Atherosclerosis 2010 Aug;211(2):539-544. [doi: 10.1016/j.atherosclerosis.2010.03.027] [Medline: 20409549]
12. Poch E, González D, Giner V, Bragulat E, Coca A, de La Sierra A. Molecular basis of salt sensitivity in human hypertension. Evaluation of renin-angiotensin-aldosterone system gene polymorphisms. Hypertension 2001 Nov;38(5):1204-1209. [Medline: 11711524]
13. Edwards KT, Huang CJ. Bridging the consumer-medical divide: how to regulate direct-to-consumer genetic testing. Hastings Cent Rep 2014 May;44(3):17-19. [doi: 10.1002/hast.310] [Medline: 24821248]
14. Covolo L, Rubinelli S, Ceretti E, Gelatti U. Internet-based direct-to-consumer genetic testing: a systematic review. J Med Internet Res 2015 Dec 14;17(12):e279 [FREE Full text] [doi: 10.2196/jmir.4378] [Medline: 26677835]
15. Goldsmith L, Jackson L, O'Connor A, Skirton H. Direct-to-consumer genomic testing from the perspective of the health professional: a systematic review of the literature. J Community Genet 2013 Apr;4(2):169-180 [FREE Full text] [doi: 10.1007/s12687-012-0135-8] [Medline: 23322235]
16. Nielsen DE, El-Sohemy A. Disclosure of genetic information and change in dietary intake: a randomized controlled trial. PLoS One 2014 Nov;9(11):e112665 [FREE Full text] [doi: 10.1371/journal.pone.0112665] [Medline: 25398084]
17. Livingstone KM, Celis-Morales C, Navas-Carretero S, San-Cristobal R, Macready AL, Fallaize R, Food4Me Study. Effect of an internet-based, personalized nutrition randomized trial on dietary changes associated with the Mediterranean diet: the Food4Me Study. Am J Clin Nutr 2016 Aug;104(2):288-297 [FREE Full text] [doi: 10.3945/ajcn.115.129049] [Medline: 27357094]
18. Fallaize R, Celis-Morales C, Macready AL, Marsaux CF, Forster H, O'Donovan C, Food4Me Study. The effect of the apolipoprotein E genotype on response to personalized dietary advice intervention: findings from the Food4Me randomized controlled trial. Am J Clin Nutr 2016 Sep;104(3):827-836. [doi: 10.3945/ajcn.116.135012] [Medline: 27510539]
19. O'Donovan CB, Walsh MC, Gibney MJ, Brennan L, Gibney ER. Knowing your genes: does this impact behaviour change? Proc Nutr Soc 2017 Dec;76(3):182-191. [doi: 10.1017/S0029665116002949] [Medline: 28103960]
20. Celis-Morales C, Marsaux CF, Livingstone KM, Navas-Carretero S, San-Cristobal R, Fallaize R, et al. Can genetic-based advice help you lose weight? Findings from the Food4Me European randomized controlled trial. Am J Clin Nutr 2017 May;105(5):1204-1213. [doi: 10.3945/ajcn.116.145680] [Medline: 28381478]
21. Littlejohn P, Cop I, Brown E, Afroze R, Davison KM. Comparison of nutrigenomics technology interface tools for consumers and health professionals: protocol for a mixed-methods study. JMIR Res Protoc 2018 Jun 11;7(6):e115 [FREE Full text] [doi: 10.2196/resprot.9846] [Medline: 29891470]
22. Institute of Medicine. Dietary Reference Intakes: Applications in Dietary Assessment. Washington, USA: National Academies Press; 2001.
23. Health Canada. Eating Well With Canada's Food Guide. Ottawa, Canada: Health Canada; 2007.
24. British Columbia Government. British Columbia Nutrition Survey - Report on Food Group Use. Victoria, Canada: BC Ministry of Health Services; 2004.
25. Grimaldi KA, van Ommen B, Ordovas JM, Parnell LD, Mathers JC, Bendik I, et al. Proposed guidelines to evaluate scientific validity and evidence for genotype-based dietary advice. Genes Nutr 2017 Dec;12:35 [FREE Full text] [doi: 10.1186/s12263-017-0584-0] [Medline: 29270237]
26. SurveyMonkey Inc. 2017. URL: http://www.surveymonkey.com
27. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
28. Health Canada. BC Nutrition Survey Training Manual. Ottawa, Canada: Health Canada; 1999.
29. Garriguet D. Diet quality in Canada. Health Rep 2009 Sep;20(3):41-52 [FREE Full text] [Medline: 19813438]
30. Lowe MR. The effects of dieting on eating behavior: a three-factor model. Psychol Bull 1993 Jul;114(1):100-121. [Medline: 8346324]
31. Turner-Bowker DM, Bayliss MS, Ware JE, Kosinski M. Usefulness of the SF-8 Health Survey for comparing the impact of migraine and other conditions. Qual Life Res 2003 Dec;12(8):1003-1012. [Medline: 14651418]
32. Ware J, Kosinski M, Dewey J, Gandek B. How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the SF-8 Health Survey. Boston, USA: QualyMetric; 2001.
33. Schwarzer R, Jerusalem M. Generalized Self-Efficacy Scale. In: Weinman J, Wright S, Johnston M, editors. Measures in health psychology: A User's Portfolio. Causal and Control Beliefs. Windsor, England: NFER-NELSON; 1995.
34. Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005 Sep;139(5):439-457. [doi: 10.3200/JRLP.139.5.439-457] [Medline: 16285214]
35. Prochaska JO, DiClemente CC. Stages of change in the modification of problem behaviors. Prog Behav Modif 1992;28:183-218. [Medline: 1620663]
36. Taylor HL, Jacobs DR, Schucker B, Knudsen J, Leon AS, Debacker G. A questionnaire for the assessment of leisure time physical activities. J Chronic Dis 1978;31(12):741-755. [Medline: 748370]
37. Statistics Canada. Canadian Community Health Survey. Ottawa, Canada: Statistics Canada; 2010. URL: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=81424

XSL•FO
RenderX

38. Yu L, Buysse DJ, Germain A, Moul DE, Stover A, Dodds NE, et al. Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. Behav Sleep Med 2011 Dec 28;10(1):6-24 [FREE Full text] [doi: 10.1080/15402002.2012.636266] [Medline: 22250775]
39. ESHA Research: Food Labeling Software | Nutrition Analysis Software. Salem, OR: USAHA Research; 2017. Food Processor SQL URL: https://www.esha.com/
40. Health Canada. Canadian Nutrient File. Ottawa, Canada: Health Canada; 2016.
41. Stata Corporation. Stata Statistical Software: Release 2. College Station, USA: Stata Corporation; 2011.
42. NVivo qualitative data analysis software, Version 11. Victoria, Australia: QSR International Pty Ltd; 2015. URL: https://www.qsrinternational.com/nvivo/nvivo-products
43. Sandelowski M. Qualitative analysis: what it is and how to begin. Res Nurs Health 1995 Aug;18(4):371-375. [Medline: 7624531]
44. Forster H, Walsh MC, O'Donovan CB, Woolhead C, McGirr C, Daly EJ, et al. A dietary feedback system for the delivery of consistent personalized dietary advice in the web-based multicenter Food4Me study. J Med Internet Res 2016 Jun 30;18(6):e150 [FREE Full text] [doi: 10.2196/jmir.5620] [Medline: 27363307]
45. Marteau TM, French DP, Griffin SJ, Prevost AT, Sutton S, Watkinson C, et al. Effects of communicating DNA-based disease risk estimates on risk-reducing behaviours. Cochrane Database Syst Rev 2010 Oct 6(10):CD007275 [FREE Full text] [doi: 10.1002/14651858.CD007275.pub2] [Medline: 20927756]
46. Rankin A, Bunting BP, Poínhos R, van der Lans IA, Fischer AR, Kuznesof S, et al. Food choice motives, attitude towards and intention to adopt personalised nutrition. Public Health Nutr 2018 Oct;21(14):2606-2616. [doi: 10.1017/S1368980018001234] [Medline: 29771231]
47. O'Donovan CB, Walsh MC, Forster H, Woolhead C, Celis-Morales C, Fallaize R, et al. The impact of 677C → T risk knowledge on changes in folate intake: findings from the Food4Me study. Genes Nutr 2016;11:25 [FREE Full text] [doi: 10.1186/s12263-016-0539-x] [Medline: 27708721]
48. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. Lancet 2017 Dec 7;389(10064):37-55 [FREE Full text] [doi: 10.1016/S0140-6736(16)31919-5] [Medline: 27863813]
49. Global Health Observatory (GHO) data. Geneva, Switzerland: World Health Organization; 2015. URL: http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence/en
50. Jing X, Chen J, Dong Y, Han D, Zhao H, Wang X, et al. Related factors of quality of life of type 2 diabetes patients: a systematic review and meta-analysis. Health Qual Life Outcomes 2018 Sep 19;16(1):189 [FREE Full text] [doi: 10.1186/s12955-018-1021-9] [Medline: 30231882]

Abbreviations

C: control group
DFE: dietary folate equivalent
FFQ: Food Frequency Questionnaire
GSE: General Self-Efficacy
I: intervention group
HEI-C: Healthy Eating Index–Canadian
HRQoL: health-related quality of life
MD: mean difference
NHP: natural health product
RD: registered dietitian

Edited by G Eysenbach; submitted 23.10.18; peer-reviewed by A Rankin, E Seto; comments to author 21.03.19; revised version received 05.05.19; accepted 05.05.19; published 28.06.19.

Please cite as:
Araujo Almeida V, Littlejohn P, Cop I, Brown E, Afroze R, Davison KM
Comparison of Nutrigenomics Technology Interface Tools for Consumers and Health Professionals: A Sequential Explanatory Mixed Methods Investigation
J Med Internet Res 2019;21(6):e12580
URL: http://www.jmir.org/2019/6/e12580/
doi:10.2196/12580
PMID:31254340

XSL•FO
RenderX

©Vanessa Araujo Almeida, Paula Littlejohn, Irene Cop, Erin Brown, Rimi Afroze, Karen M Davison. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Promotion of Vape Tricks on YouTube: Content Analysis

Grace Kong[1], PhD; Heather LaVallee[1], MA, MFT; Alissa Rams[1], BA; Divya Ramamurthi[2], MA; Suchitra Krishnan-Sarin[1], PhD

[1]Yale School of Medicine, New Haven, CT, United States
[2]Stanford University, Stanford, CA, United States

Corresponding Author:
Grace Kong, PhD
Yale School of Medicine
34 Park St
New Haven, CT, 06519
United States
Phone: 1 2039747601
Email: grace.kong@yale.edu

Abstract

Background: The ability to perform vape tricks (ie, blowing large vapor clouds or shapes like rings) using e-cigarettes appeals to youth. Vape tricks are promoted on social media, but the promotion of vape tricks on social media is not well understood.

Objective: The aim of this study was to examine how vape tricks were promoted on YouTube to youth.

Methods: Videos on vape tricks that could be accessed by underage youth were identified. The videos were coded for number of views, likes, dislikes, and content (ie, description of vape tricks, e-cigarette devices used for this purpose, video sponsors [private or industry], brand marketing, and contextual characteristics [eg, model characteristics, music, and profanity]).

Results: An analysis of 59 sample videos on vape tricks identified 25 distinct vape tricks. These videos had more likes than dislikes (11 to 1 ratio) and a 32,017 median view count. 48% (28/59) of the videos were posted by industry accounts (27% [16/59] provaping organizations, 15% [9/59] online shops, and 3% [2/59] vape shops) and 53% by private accounts (55% [17/31] private users, 26% [8/31] vape enthusiasts, and 19% [6/31] YouTube influencers); 53% (31/59) of the videos promoted a brand of e-cigarette devices, e-liquids, or online/vape shops, and 99% of the devices used for vape tricks were advanced generation devices. The models in the videos were 80.2% (160/198) male, 51.5% white (102/198), and 61.6% (122/198) aged 18 to 24 years; 85% (50/59) of the videos had electronic dance music and hip hop, and 32% (19/59) had profanity.

Conclusions: Vape trick videos on YouTube, about half of which were industry sponsored, were accessible to youth. Restrictions of e-cigarette marketing on social media, such as YouTube, are needed.

(J Med Internet Res 2019;21(6):e12709) doi:10.2196/12709

KEYWORDS

e-cigarettes; social media; marketing

Introduction

Background

E-cigarettes have rapidly gained popularity among adolescents and adults in the United States and around the world since their introduction to the market around 2007. US national data show that although current e-cigarette use (defined as using it every day or some days) is 3.3% among adults older than 24 years, use among young adults aged 18 to 24 years is higher at 5.2% [1], and current use (defined as using it one day or more in the past 30 days) is 20.8% among high school students [2].

Understanding the appeal of e-cigarettes is critical to inform e-cigarette prevention efforts aimed at adolescents and young adults. Vape tricks, which involve using e-cigarettes to blow large, thick amounts of exhaled aerosol (ie, clouds) or shapes such as rings, are known to be appealing to youth [3,4]. There are even competitions held with prizes, sponsors, judges, and spectators to see who can produce impressive vape tricks using e-cigarettes [5]. The appeal of vape tricks appears to be unique to youth as this appeal has been identified as a reason for e-cigarette initiation among youth e-cigarette users [3,4] but not among adult users [6]. However, despite the appeal of vape tricks among youth, there is limited information on this e-cigarette use behavior. It is important to understand how vape

XSL•FO
RenderX

tricks are conducted and promoted to youth as the appeal of vape tricks could lead youth to initiate e-cigarettes [3,4] and expose them to nicotine dependence and to other unknown health effects from using e-cigarettes.

Use of Social Media to Understand E-Cigarette Use

Social media provides a unique opportunity to understand health-related behaviors [7] and novel tobacco use behaviors, including e-cigarettes [8]. Social media is particularly an important medium to examine as youth are actively engaging in social media, and importantly, they are exposed to protobacco content through this medium [9]. Youth exposure to protobacco content on social media is concerning owing to its influence on youth behaviors and perceptions.

According to the Social Cognitive Theory of Mass Communication [10], exposure to social media influences youth's cognition, emotion, and behavior. Thus, positive perceptions regarding e-cigarettes and e-cigarette use behaviors, such as conducting vape tricks, may be perpetuated by e-cigarette-related contents portrayed on social media. Indeed, existing studies have examined e-cigarette-related perceptions and use behaviors on popular social media websites among youth, such as Reddit [11,12], Twitter [13-15], Instagram [16-19], and YouTube [20-22]. Thus, analysis of social media data provides an important means to understand how e-cigarette use behaviors are promoted without being primed by a researcher.

Use of YouTube to Understand Vape Tricks

In this study, we have focused on YouTube to learn how vape tricks are promoted to youth. YouTube is a free and popular website that allows users to upload and share videos and it is extremely popular among youth. YouTube has over a billion users, and more young people (aged 18 to 34 years) are watching YouTube videos than any cable network channels [23]. Moreover, adolescents who are younger than 18 years are 1.5 times more likely than adults to frequent YouTube [24]. Adolescents are also more likely than adults to be exposed to alcohol and tobacco content on YouTube videos [25].

YouTube is a useful source to learn about traditional and novel tobacco use behaviors. For instance, YouTube has been used to understand unorthodox tobacco use methods, such as techniques on manipulating cigars [26]. YouTube videos are widely used to promote various tobacco products, including cigarettes [27,28], hookah [29], smokeless tobacco [30,31], cigars and little cigars [32], and e-cigarettes [20,22,33-36]. For instance, vape pens have been promoted via music videos on YouTube, and these music videos have been viewed over 1.4 billion times in 24 months [37]. The widespread presence of protobacco content on YouTube is concerning as it can influence perceptions and behaviors of their young viewers. As YouTube is readily available to all age groups, and may have a particularly greater influence on adolescents, it is important to understand if and how this social media platform is being used to promote vape tricks. Youth may turn to YouTube to learn how to conduct new behaviors that they have never conducted before, such as vape tricks [38].

The Goal of This Study

We examined the content of YouTube videos that promote vape tricks, and which may be accessible to youth, using a nonage verified account. We used search terms related to tutorials and *how to* videos (eg, step-by-step instructions on how to conduct vape tricks) on conducting vape tricks on YouTube to assess the types of information young users may be exposed to. We aimed to describe the availability of these videos, the video characteristics (eg, number of views and *likes*), and the type of vape tricks shown and e-cigarette devices used for this purpose.

Marketing e-cigarettes on social media poses a significant challenge to tobacco control as this content remains relatively unknown. Thus, we determined the source of these videos–whether these videos were posted by the e-cigarette industry (eg, e-cigarette device/e-liquid manufacturers, vape shops, or pro–e-cigarette organizations) or by private users who have no clear links with the e-cigarette industry. We further determined whether industry-sponsored videos had more marketing content than private-user videos. Although it may be evident that industry-sponsored videos would have more marketing content than private-user videos, social media presents novel marketing opportunities for private users to also market products to the public [39-41]. Thus, we also assessed whether private users were YouTube influencers or vape enthusiasts. YouTube influencer is a celebrity who usually posts videos of themselves (eg, vlogging) on a subject matter in which they view themselves as an expert or on other various subject matters through their YouTube channels. They often have a large following and thus possess the potential to reach and market products to many viewers [41,42]. *Vape enthusiasts* are individuals who are experts in vaping and may use social media to promote vaping and market-related products.

We also described contextual video characteristics that could enhance the appeal of these videos to young viewers, such as the production quality of the videos, model characteristics (eg, gender and age), and other appealing aspects such as music. Finally, we assessed the use of profanity in the videos; as these are videos that could also be accessed by youth, the identification of the presence of profanity may be helpful to setting age restrictions to accessing these videos [43].

Methods

Procedures

In April/May 2016, we searched for terms related to vape tricks on YouTube, including tutorials/instructions on how to conduct these behaviors (ie, vape tricks, e-cig smoke tricks, how to do vape tricks, how to do smoke tricks with vape, how to do smoke tricks with vape pens, how to do smoke tricks for beginners vape, how to do smoke tricks for beginners e-cigs, vape tricks tutorial, e-cig smoke tricks tutorial, e-cigarette smoke trick tutorial, and vape smoke tricks tutorial). YouTube restricts minors younger than 18 years from videos with inappropriate contents, such as vulgar language, violence and disturbing imagery, nudity/sexually suggestive content, and portrayal of harmful or dangerous activities [44]. To access these videos, viewers must register their date of birth on their Google account. However, users not signed into an age-verified account could

XSL•FO
RenderX

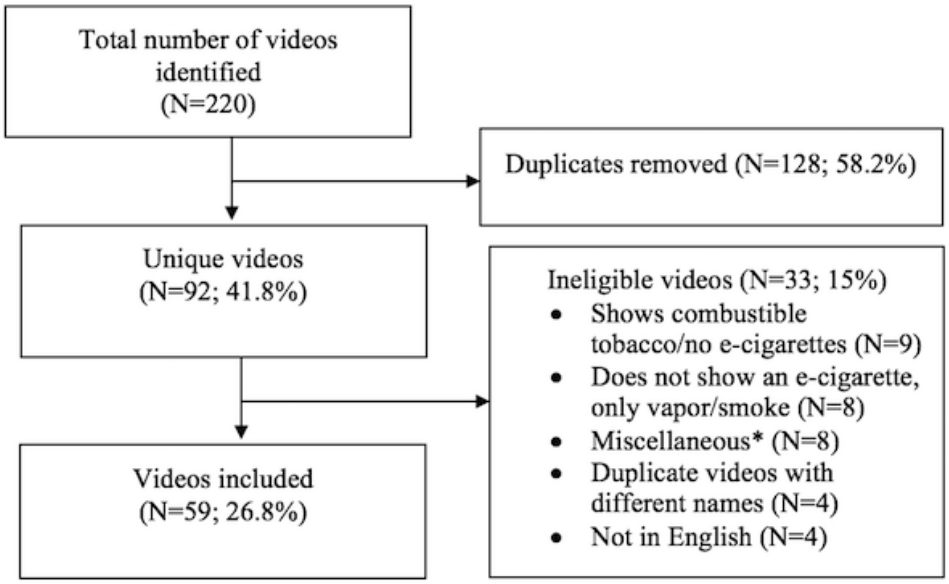
access all other videos that are not flagged as inappropriate. Adolescents may sign in from accounts that are not age verified or from age-verified accounts that automatically prohibit them from accessing flagged videos, so we searched for the terms from a YouTube account that was not age verified.

We used search parameters that have been used to examine tobacco use behaviors on YouTube [29,33,45]: (1) the search was limited to videos uploaded in the past year with a duration of <4 min, (2) the videos were sorted by relevance (default sorting option), and (3) videos on the first 2 pages (20 videos) using each of the search terms were downloaded. Previous studies showed that the majority of internet users click on the first page of search results [46]. Based on this finding, studies examining YouTube videos examined the first 20 videos shown on the first 2 pages of the search results with the assumption that users would not watch more than 20 videos [29,33,35]. To be included for analysis, videos had to be in English, show a vape trick, and an e-cigarette to do the vape trick to confirm that the vape tricks were conducted using e-cigarettes and not other tobacco products such as cigarettes or hookah. See Figure 1 for the number of videos identified and reasons for exclusion.

To analyze the videos, a codebook was developed based on a previously conducted analysis of tobacco-related content of YouTube videos [26]. Specifically, we obtained descriptive information about YouTube videos, such as the number of comments, views, likes, dislikes, and duration of the videos. Content analysis included vape trick characteristics (ie, identification and description of vape tricks and e-cigarettes used for this purpose), video source (ie, private vs e-cigarette industry and the presence of marketing content), and contextual characteristics that enhanced the appeal of the videos such as models and music. A detailed description of each content area is provided below.

The final codebook was developed iteratively. These content areas were identified and described in detail by the lead researcher who viewed videos that met the criteria except for the year of publication (ie, videos published 6 months before the specified dates used in search) and then they were confirmed or modified by 2 independent coders who also reviewed the same videos. All coders resolved any differences in coding through discussions. Upon the development of the codebook, we determined interrater reliability by coding 30% of videos not included in the final analysis (ie, videos uploaded 6 months before the specified dates used in search). To establish interrater reliability, 2 independent coders coded 17 videos. The coding was then reviewed by 2 coders and by the leader researcher to amend coding differences and answer any questions that arose during coding. The coders then coded the videos included in this study. Each video was viewed as many times as possible to code all aspects of the interested variables.

Figure 1. Inclusion of YouTube videos on vape tricks and reasons for exclusion. The videos shown on the first 2 pages of each search term (n=11) were considered for inclusion. *Miscellaneous ineligible videos included videos on magic tricks, how to hack an e-cigarette not for conducting vape tricks, and video games unrelated to vape tricks.



Content Areas

Vape Trick Characteristics

Vape tricks were identified and described based on the information provided by the models on the YouTube videos and the visual and verbal demonstrations of vape tricks shown on the videos. We described each vape trick and determined how many times each vape trick was shown on all videos. The Cohen kappa was 0.69, $P<.001$.

E-cigarette devices (eg, box mods, vape pens, and cigalikes) used for vape tricks were identified through visual confirmation by the coders. The Cohen kappa was 0.85, $P<.001$.

Video Source Characteristics

Video source was defined as an e-cigarette industry (eg, e-cigarette or e-liquid manufacturers, online shops, vape shops, or pro–e-cigarette professional organizations) or a private person who did not have a clear affiliation to an e-cigarette industry and appeared to represent only himself/herself. To determine

XSL•FO
RenderX

the video source, we examined the channel pages of each YouTube video. The channel pages referred to the profile page of the YouTube user identifier who posted the video. This page could be used to obtain information about the posters, such as their affiliated businesses or organizations, when applicable. Industry-sponsored channels were further categorized into (1) vape shops if the channel was associated with specific brick and mortar stores, (2) Web-based shops if they were selling e-cigarettes or related products online but were not associated with brick and mortar stores, or (3) pro–e-cigarette organizations if they endorsed vaping by providing reviews about e-cigarette devices, vape tricks, or shared other news associated with e-cigarettes but did not sell any tobacco products.

We also assessed whether the videos were posted by a YouTube influencer. YouTube influencer was defined as a single key person who had a significant presence on YouTube with either many subscribers (>10,000) or more than 1 million views of their channels. Another source of influence may be a *vape enthusiast*. We determined *vape enthusiast* on the basis of the description provided on the channel page and if they only included videos on e-cigarettes on their video playlist. The Cohen kappa was 0.92, *P*<.001.

Marketing was determined if there was verbal or nonverbal brand promotion of e-cigarette devices, e-liquids, vape shops, or e-cigarette–related organizations. Examples of verbal marketing were a model saying the brand names of the devices or the e-liquids that they were using to conduct vape tricks or acknowledging their sponsors (eg, vape shops) on the video. Examples of nonverbal brand marketing included displaying the brand names or the images of the logo of the devices or the e-liquids or sponsor on the screen, on clothing or hat worn by the models, or on signs and posters in the background. Cohen kappa was 0.63, *P*<.001.

Contextual Characteristics

Video type was categorized into (1) tutorials: step-by-step instructions on how to conduct vape tricks, (2) compilations: amalgam of video clips of various types of vape tricks in 1 video, and (3) *other*: other types of video that could not be categorized into the specified categories. Cohen kappa was 0.84, *P*<.001.

Production quality was determined by ratings of low, moderate, or high based on previously used coding of YouTube videos [29]. The videos were coded low if they looked like they were homemade and there was little or no attention to production values such as lighting, camera angles, and sound quality; moderate if they were homemade but at least some attention was paid to production values; high if the videos were produced with great attention to these production values. Cohen kappa was 0.68, *P*<.001.

Model characteristics such as gender, race, and age were determined for each model shown on the video using the coding scheme previously used to code models on smoking videos on YouTube [45]. Gender was coded as male, female, and cannot identify. Perceived race was coded as white (including European), Asian (including Asian-American), black/African American, Latino, and cannot Identify. Perceived age included younger than 18 years, 18 to 24 years, 25 to 34 years, 35 to 59 years, 60 years or older, and cannot identify. Cohen kappa for each demographic variable was 0.92, 0.63, 0.83, *P*<.001, respectively.

Other contextual elements such as the use of music and profanity were assessed. We coded the music into categories of electronic dance music (EDM), hip hop, pop, and classical. We also coded for the use of verbal and nonverbal profanity (ie, sticking out the middle finger). Cohen kappa for music was 1.00, *P*<.001, and Cohen kappa for profanity was 0.74, *P*<.001.

Results

Video Characteristics

The 11 search terms identified a total of 156,200 videos (mean 14,200 [SD 12,529]; range 2190-43,100). The videos shown only on the first 2 pages (20 videos on each page) for each search term (N=220) were considered. After the removal of duplicate videos (n=128) and the videos that did not meet our inclusion criteria (n=33), 59 videos were included in the final analysis (see Figure 1 for specific reasons for exclusion). Table 1 lists the descriptive statistics of the videos, such as the published date of the videos, duration of the videos, and the number of views, likes, dislikes, and comments. The average length/duration of the videos was 2 min 14 s (SD 55 seconds), and the videos had more *likes* than *dislikes* (11 to 1 ratio).

Vape Trick Characteristics

Vape Tricks

We identified 25 distinct vape tricks. The median number of vape tricks shown on each video was 5 (interquartile range=4, range 1-20). See Figure 2 for the names of the vape tricks and the frequencies of vape tricks shown on the videos; a further description of these vape tricks was provided in the supplemental materials. The most common vape trick was a *modified trick*, which was a modification of any existing vape trick.

E-Cigarette Devices Used for Vape Tricks

We were able to determine the type of e-cigarette devices for 72.0% (136/189) of the devices shown . A total of 98.5% (134/136) of the devices were advanced generation devices (44.8% box mod (61/136) and 53.7% (73/136) vape pens) and 1.5% (2/136) of the devices were cigalikes. Box mods and vape pens were further categorized into whether they were drip or tank devices (Figure 3).

XSL•FO
RenderX

Table 1. Descriptive statistics of vape trick video characteristics.

Video characteristics	Median	Interquartile range	Minimum value	Maximum value
Published date	November 20, 2016	200 days	April 29, 2015	April 13, 2016
Views	32,017	122,842	44	2,331,038
Likes	194	596	1	7439
Dislikes	8	35	0	1787
Duration (Min:Seconds)	2:08	1:30	0:17	3:52
Comments	21	50	0	1384

Figure 2. The 25 vape tricks identified in YouTube videos. See Supplemental material for the description of these vape tricks. The frequencies were based on the number of times these vape tricks appeared on all videos. Each video may have featured more than one vape trick.

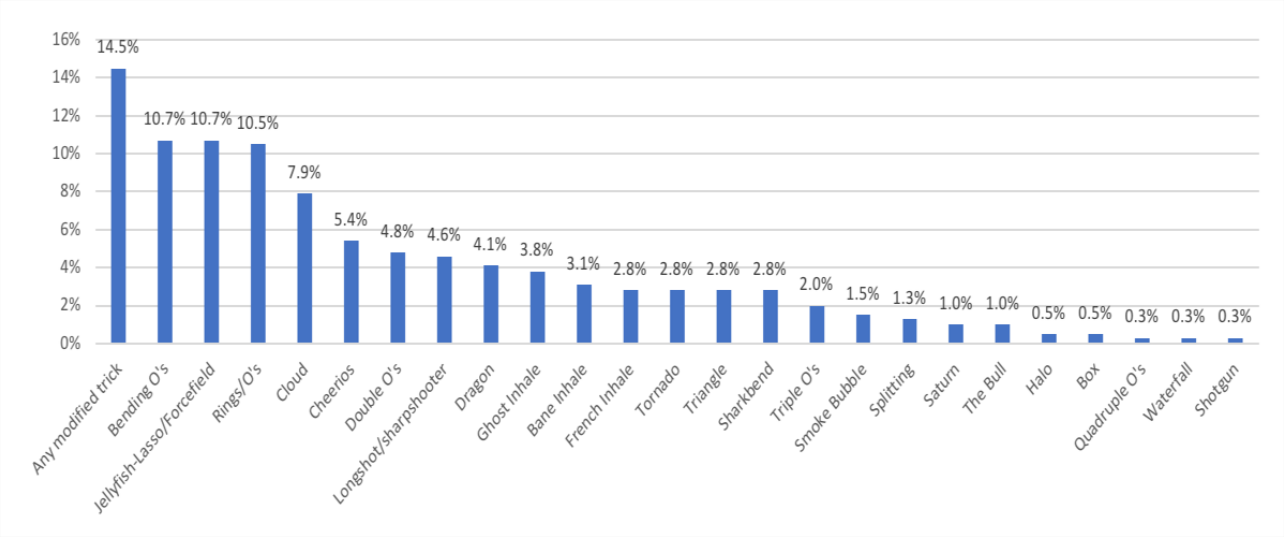
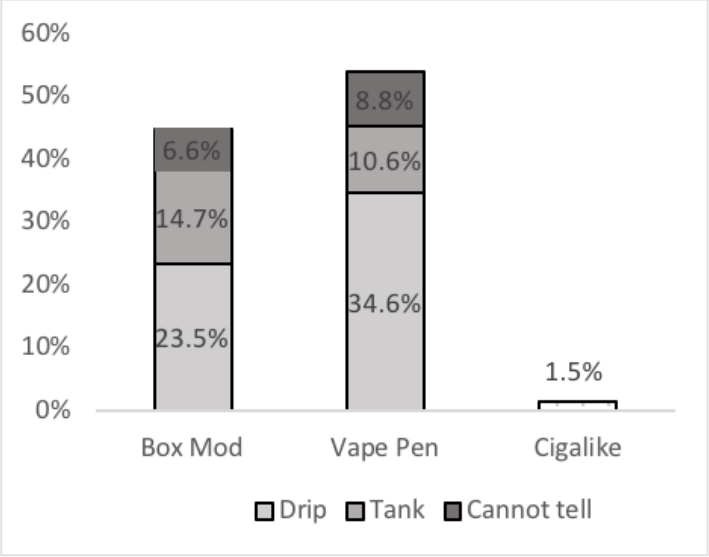


Figure 3. E-cigarette devices used for vape tricks on YouTube videos.



Video Source Characteristics

Video Source

Personal channels posted 53% (31/59) of the videos and industry-sponsored channels posted 48% (28/59) of the videos (27% (16/59) provaping organizations, 15.3% (9/59) Web-based shops, and 3.4% (2/59) vape shops). Of the private channels, 55% (17/31) were by private user channels, 26% (8/31) by vape enthusiasts, and 19% (6/31) by YouTube influencers. Interestingly, only 1 influencer had videos that showcased only e-cigarette–specific video content whereas the other influencers posted videos on various content areas (eg, tattoo, music videos, humor, and viral videos) including content on e-cigarettes. Of the private channels, 25.8% were by "vape enthusiasts," and 54.8% were private-user channels.

XSL•FO
RenderX

Marketing

53% (31/59) of the videos contained marketing content (25% [15/59] nonverbal only, 5% [3/59] verbal only, and 22% [13/59] both nonverbal and verbal). Some examples of nonverbal brand promotion were models conducting vape tricks while wearing shirts or hats with the brand logo, showing an e-liquid bottle with the brand logo and name, or the brand name was displayed on the screen. An example of verbal promotion included the model verbally encouraging the viewers to purchase the e-liquid or devices.

Extraneous Video Characteristics

Video Type

Video types are as follows: 41% (24/59) of the videos were tutorial, 34% (20/59) were compilations, and 25% (15/59) were other. Other videos included short clips of vape trick competitions, model(s) conducting vape tricks on a stage with special lighting and music, and selfie videos of individuals conducting vape tricks.

Production Quality

Production quality are as follows: 25% (15/59) were low, 34% (20/59) were moderate, and 41% (24/59) were high quality. Videos sponsored by the industry were more likely to be of high quality relative to personal videos (personal videos: low quality 42% (13/31), moderate quality 52% (16/31), and high quality 7% (2/31) vs industry videos: low quality 7% (2/28), moderate quality 14% (4/28), and high quality 79% (22/28); X^2_2=31.86; P<.001).

Model Characteristics

There were 198 models in total: 80.2% (160/198) were male, 18.2% (36/198) were female, and 1.0% (2/198) could not be identified. Race varied: 51.5% (102/198) were coded as white, 17.7% (35/198) as Asian, 17.2% (34/198) as Latino, 3.0% (6/198) as black, and 10.6% (21/198) could not be identified. The age groups were 61.6% (122/198) 18 to 24 years, 26.3% (52/198) 25 to 34 years, 5.1% (10/198) younger than 18 years, 1.0% (2/198) older than 35 years, and 6.1% (12/198) could not be identified.

Other Contextual Elements

Music

A total of 58% (34/59) of the videos had only EDM, 15% (9/59) had EDM and other music, such as hip hop and pop, 14% (8/59) of the videos did not have any music, 12% (7/59) had only hip hop, and 2% (1/59) had classical music.

Profanity

A total of 32% (19/59) of the videos contained profanity: 10% (6/59) had verbal profanity, 9% (5/59) had models using nonverbal and visual profanity (ie, sticking out the middle finger), 9% (5/59) had music containing profanity, and 5% (3/59) had multiple sources of profanity (usually a mix of music and verbal profanity).

Discussion

Principal Findings

The search of vape tricks and tutorials on vape tricks on YouTube resulted in an average of 14,200 YouTube videos, which were videos that could be accessed by underage youth (<18 years). This finding could serve as an impetus for YouTube to set limits to restrict underage youth from accessing videos on vape tricks. YouTube self-regulates by attempting to limit youth from accessing inappropriate videos that are deemed sexually explicit, harmful or dangerous, violent or graphic, hateful, spams/scams, threats, excessive profanity, and copyrighted [47]. To access these flagged videos, users must use age-verified accounts. Currently, protobacco videos do not fall under the definition of unacceptable materials and are readily available without any regulations. Given that the age of legal tobacco use is 18 years and older (in some places 21 years and older), YouTube may consider videos on e-cigarettes and other tobacco use as mature content and flag these videos so that users with age-verified accounts can access them. Furthermore, one-third of the videos in our study contained profanity, which further suggests that these videos are not suitable for young viewers.

The evaluation of sample videos on vape tricks identified 25 distinct vape tricks, which were mostly performed by young adults using advanced generation e-cigarette devices. Advanced generation devices are rechargeable and are highly customizable; they include vape pens, box mods, and mech-mods. In addition, some of these devices have specific features for dripping (ie, applying drops of e-liquid solution onto an atomizer to saturate the wick before coil heating) [48]. The promotion of advanced generation devices, including dripping to conduct vape tricks on YouTube, is concerning as studies have shown that adolescent and young adult e-cigarette users primarily use advanced generation e-cigarettes [49,50], whereas adult users use cigalikes [51]. Moreover, data from Connecticut showed that 26% of adolescent e-cigarette users are dripping [50]. Advanced generation e-cigarette devices may be popular among youth because of the ability to customize the product to drip and also to change the temperature of the devices and the constituents of the e-liquid, such as the flavors, nicotine level, and the ratio of the solvent (propylene glycol [PG] and vegetable glycerin [VG]). These components could be modified to conduct vape tricks. For example, temperatures of the devices and PG/VG ratio of the e-liquids could be modified to produce a large amount of exhaled aerosol to conduct vape tricks.

Conducting vape tricks using advanced generation devices is problematic as the use of these devices can be harmful to users because the increased level of temperature and the higher PG levels are associated with a greater exposure to nicotine and toxicants [52,53]. Furthermore, individuals engaging in vape tricks may alter how they puff by taking in deeper and more frequent puffs, which may also increase health risks. Another concern is the exposure to nicotine dependence and other tobacco use behaviors by conducting vape tricks. For example, youth may be drawn to e-cigarettes via YouTube videos displaying appealing vape tricks and they may practice and use

XSL•FO
RenderX

e-cigarettes more frequently to learn how to conduct certain vape tricks, which may expose them to more nicotine and other chemicals in the e-cigarette. However, the vape trick videos on YouTube evaluated in this study did not have information on how to manipulate the devices or to change puffing behaviors to conduct various vape tricks. Future studies should assess which and how e-cigarettes are being customized to conduct vape tricks and if puffing behaviors are altered to assess potential health risks for engaging in this behavior.

We also observed that about half of the YouTube videos that promoted vape tricks were sponsored by the e-cigarette industry, with pro–e-cigarette organizations (which are organizations that promote and advocate for pro–e-cigarette–related issues and policies [54]) being the most common sponsor, followed by online shops and vape shops. These findings suggest that when trying to understand the e-cigarette industry, it is important to recognize that stakeholders such as pro–e-cigarette organizations, online retailers, and vape shops are actively engaging in marketing and promoting e-cigarettes. It is important to note that online shops were coded if they did not have a brick and mortar store associated with the shop. However, it is likely that other organizations and brick and mortar shops had online purchasing options.

The large presence of sponsored videos is consistent with the emerging literature showing that Web 2.0 is the optimal platform for the industry to market their tobacco products because of the ability to reach large number of people at fast speed and at low cost [40]. Promoting vape tricks may be a marketing strategy used by the e-cigarette industry. Vape trick videos are used to promote brands of e-liquids, devices, and vape shops, both verbally and nonverbally. Nonverbal promotion is an inconspicuous way to market as the brands are shown on the models' clothing, background, devices, or on the screen while the model is conducting vape tricks, and this subtle strategy is an effective marketing strategy [55]. Although verbal and nonverbal promotion is one clear way in which e-cigarettes are marketed, future studies should examine a greater number of videos to understand nuanced marketing strategies used in both industry-sponsored and private videos, using both verbal and nonverbal strategies.

One novel marketing strategy used on social media is the use of YouTube influencers or vaping enthusiasts to promote e-cigarette use and market-associated tobacco products. Indeed, we found that vape trick videos were uploaded by these private users. We also observed that 39% of the videos posted by private users (which includes YouTube influencers and vape enthusiasts) also promoted a specific brand of e-liquid or e-cigarette devices. As e-cigarette marketing on social media continues to grow and evolve, more research is needed on the marketing strategies used on social media platforms to inform restrictions on tobacco marketing.

In addition to the type of vape tricks being conducted, and the devices being used for this purpose, the videos had contextual factors that could be appealing to youth. Identification of such factors may reveal how the appeal of vape tricks is enhanced. Despite using various terms related to tutorials on vape tricks, we found that 59% of the videos did not contain information on how to conduct vape tricks but showcased vape tricks in various ways, such as a compilation of videos clips of vape tricks into 1 video accompanied by music, similar to a music video; individual(s) performing vape tricks on a stage accompanied by special lighting and music; vape trick competition videos, which were mostly brief summaries of vape trick competitions featuring winners and their featured vape trick and *news reporters* interviewing the winners; and *selfie* videos where a user self-records himself/herself conducting vape tricks. These various methods of showcasing vape tricks could enhance the appeal of e-cigarettes to youth.

Interestingly, our examination of YouTube videos shows that vape tricks are being conducted as competitions. Future studies are needed to assess the prevalence of these competitions, who attends them, who promotes them, and how these competitions are used to promote e-cigarettes.

The identified contextual factors of YouTube videos may contribute to the coolness of conducting vape tricks. Research on perceptions has shown that short-term benefits such as looking cool are associated with tobacco use and intent to use tobacco [56]. Even though the self-report data suggest that vape tricks are appealing to adolescents [3,4], data on youth's perceptions of vape trick videos on YouTube are lacking. However, the high number of views and likes does suggest popularity. In fact, the view count for vape trick videos identified in this study was comparable with previous research on videos on e-cigarettes [22] and smokeless tobacco [30] identified on YouTube.

It is interesting to note that 80% of the models in the videos were males. Although national data have shown that e-cigarette use is greater among male adolescents and adults [57,58], whether vape tricks are more popular among males is unknown. Although it is possible that e-cigarette use and associated behaviors such as vape tricks are more common among males, it is also possible that males are more likely to post videos of themselves conducting vape tricks than females. Sex difference in e-cigarette use and conducting vape tricks using e-cigarettes need further investigation.

Limitations

There are several study limitations. First, we used terms related to vape trick tutorials to assess basic videos that novel users may find on YouTube. There are other terms used to refer to vape tricks, such as cloud chasing or plume tricks that were not used in this study, so the vape tricks identified in this study are not an exhaustive list. Second, we only examined YouTube and did not examine other social media platforms. Other social media platforms such as Facebook, Instagram, Twitter, and Reddit can be used to promote vape tricks and need to be examined, as they have been used to relay information about e-cigarettes [16,59]. Third, we cannot determine whether conducting vape tricks increases the risk for developing nicotine dependence and engaging in other problematic tobacco use outcomes. Future studies need to examine the role of vape tricks on initiation and progression of e-cigarette use behaviors among youth. Youth may be experimenting with vape tricks and eventually stop once the novelty wears off or this behavior may lead to continual future use of the product and expose users to

XSL•FO
RenderX

harm and other tobacco use. Fourth, we cannot determine whether youth are exposed to greater health risks by engaging in vape tricks through our study. Although our study findings did show that 99% of the vape tricks were conducted using advanced generation devices, future studies also need to examine whether youth engaging in vape tricks are engaging in riskier e-cigarette use behaviors (eg, dripping). Fifth, the kappa scores ranged from moderate to almost perfect [60]. Although these kappa values are acceptable, we did not conduct another interrater reliability after the discussion to amend any discrepancies. Future studies should conduct interrater reliability before and after the discussion to ensure that the coders are coding consistently. Finally, we did not assess how the contextual features in these vape trick videos appeal to real youth; future studies should examine which aspect of conducting vape tricks and video contextual features that are identified appeals to youth to directly assess the impact of these videos on enhancing the appeal of e-cigarettes.

Conclusions

This study also suggests areas for future e-cigarette prevention and education efforts. If YouTube videos are used to teach users about vape tricks and other e-cigarette–related information, then YouTube may be also used to disseminate health and risk information about e-cigarettes.

Acknowledgments

Research reported in this publication was supported by grant number R03DA041853 and P50DA036151 from the National Institutes of Health (NIH) National Institute on Drug Abuse (NIDA) and US Food and Drug Administration (FDA) Center for Tobacco Products. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH, NIDA, or FDA.

Conflicts of Interest

None declared.

References

1. Wang TW, Asman K, Gentzke AS, Cullen KA, Holder-Hayes E, Reyes-Guzman C, et al. Tobacco product use among adults - United States, 2017. MMWR Morb Mortal Wkly Rep 2018 Nov 9;67(44):1225-1232 [FREE Full text] [doi: 10.15585/mmwr.mm6744a2] [Medline: 30408019]
2. Gentzke AS, Creamer M, Cullen KA, Ambrose BK, Willis G, Jamal A, et al. Vital signs: tobacco product use among middle and high school students—United States, 2011-2018. MMWR Morb Mortal Wkly Rep 2019 Feb 15;68(6):157-164 [FREE Full text] [doi: 10.15585/mmwr.mm6806e1] [Medline: 30763302]
3. Kong G, Morean ME, Cavallo DA, Camenga DR, Krishnan-Sarin S. Reasons for electronic cigarette experimentation and discontinuation among adolescents and young adults. Nicotine Tob Res 2015 Jul;17(7):847-854 [FREE Full text] [doi: 10.1093/ntr/ntu257] [Medline: 25481917]
4. Pepper JK, Lee YO, Watson KA, Kim AE, Nonnemaker JM, Farrelly MC. Risk Factors for Youth E-Cigarette. J Adolesc Health 2017 Nov;61(5):599-605. [doi: 10.1016/j.jadohealth.2017.05.010] [Medline: 28712592]
5. Mickle T. The Wall Street Journal. 2015 Apr 25. Taking a deep breath if you want to try competitive vaping URL: https://www.wsj.com/articles/take-a-deep-breath-if-you-want-to-try-competitive-vaping-1429646394
6. Pepper JK, Ribisl KM, Emery SL, Brewer NT. Reasons for starting and stopping electronic cigarette use. Int J Environ Res Public Health 2014;11(10):10345-10361 [FREE Full text] [doi: 10.3390/ijerph111010345] [Medline: 25286168]
7. 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]
8. Emery SL, Vera L, Huang J, Szczypka G. Wanna know about vaping? Patterns of message exposure, seeking and sharing information about e-cigarettes across media platforms. Tob Control 2014 Jul;23(Suppl 3):iii17-iii25 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051648] [Medline: 24935893]
9. Jenssen BP, Klein JD, Salazar LF, Daluga NA, DiClemente RJ. Exposure to tobacco on the internet: content analysis of adolescents' internet use. Pediatrics 2009 Aug;124(2):e180-e186 [FREE Full text] [doi: 10.1542/peds.2008-3838] [Medline: 19620193]
10. Bandura A. Social Cognitive Theory of Mass Communication. Media Psychol 2001 Aug;3(3):265-299. [doi: 10.1207/S1532785XMEP0303_03]
11. Brett EI, Stevens EM, Wagener TL, Leavens EL, Morgan TL, Cotton WD, et al. A content analysis of JUUL discussions on social media: using Reddit to understand patterns and perceptions of JUUL use. Drug Alcohol Depend 2019 Jan 1;194:358-362. [doi: 10.1016/j.drugalcdep.2018.10.014] [Medline: 30472576]
12. Wang L, Zhan Y, Li Q, Zeng DD, Leischow SJ, Okamoto J. An examination of electronic cigarette content on social media: analysis of e-cigarette flavor content on Reddit. Int J Environ Res Public Health 2015 Nov;12(11):14916-14935 [FREE Full text] [doi: 10.3390/ijerph121114916] [Medline: 26610541]

XSL•FO
RenderX

13. Lazard AJ, Saffer AJ, Wilcox GB, Chung AD, Mackert MS, Bernhardt JM. E-cigarette social media messages: a text mining analysis of marketing and consumer conversations on Twitter. JMIR Public Health Surveill 2016 Dec 12;2(2):e171 [FREE Full text] [doi: 10.2196/publichealth.6551] [Medline: 27956376]
14. Kim AE, Hopper T, Simpson S, Nonnemaker J, Lieberman AJ, Hansen H, et al. Using Twitter data to gain insights into e-cigarette marketing and locations of use: an infoveillance study. J Med Internet Res 2015 Nov 6;17(11):e251 [FREE Full text] [doi: 10.2196/jmir.4466] [Medline: 26545927]
15. Ayers JW, Leas EC, Allem J, Benton A, Dredze M, Althouse BM, et al. Why do people use electronic nicotine delivery systems (electronic cigarettes)? A content analysis of Twitter, 2012-2015. PLoS One 2017;12(3):e0170702 [FREE Full text] [doi: 10.1371/journal.pone.0170702] [Medline: 28248987]
16. Chu KH, Allem JP, Cruz TB, Unger JB. Vaping on Instagram: cloud chasing, hand checks and product placement. Tob Control 2016 Dec;26(5):575-578 [FREE Full text] [doi: 10.1136/tobaccocontrol-2016-053052] [Medline: 27660111]
17. Phua J, Jin SV, Hahm JM. Celebrity-endorsed e-cigarette brand Instagram advertisements: effects on young adults' attitudes towards e-cigarettes and smoking intentions. J Health Psychol 2018 Dec;23(4):550-560. [doi: 10.1177/1359105317693912] [Medline: 28810409]
18. Laestadius LI, Wahl MM, Pokhrel P, Cho YI. From Apple to Werewolf: a content analysis of marketing for e-liquids on Instagram. Addict Behav 2019 Apr;91:119-127. [doi: 10.1016/j.addbeh.2018.09.008] [Medline: 30253933]
19. Laestadius LI, Wahl MM, Cho YI. #Vapelife: an exploratory study of electronic cigarette use and promotion on Instagram. Subst Use Misuse 2016 Dec 14;51(12):1669-1673. [doi: 10.1080/10826084.2016.1188958] [Medline: 27484191]
20. Sears CG, Walker KL, Hart JL, Lee AS, Siu A, Smith C. Clean, cheap, convenient: promotion of Electronic cigarettes on YouTube. Tob Prev Cessat 2017 Apr;3:10 [FREE Full text] [doi: 10.18332/tpc/69393] [Medline: 28725876]
21. Guy MC, Helt J, Palafox S, Green K, Soule EK, Maloney SF, et al. Orthodox and unorthodox uses of electronic cigarettes: a surveillance of YouTube video content. Nicotine Tob Res 2018 Jun 28:nty132. [doi: 10.1093/ntr/nty132] [Medline: 29961828]
22. Huang J, Kornfield R, Emery SL. 100 million views of electronic cigarette YouTube videos and counting: quantification, content evaluation, and engagement levels of videos. J Med Internet Res 2016 Mar 18;18(3):e67 [FREE Full text] [doi: 10.2196/jmir.4265] [Medline: 26993213]
23. YouTube. Statistics URL: https://www.youtube.com/yt/about/press/ [accessed 2019-05-22] [WebCite Cache ID 78Z0r9jz0]
24. Nielsen/NetRatings. YouTube U.S. web traffic grows 75 percent week over week, according to nielsen/netratings URL: http://www.nielsen-online.com/pr/pr_060721_2.pdf [accessed 2019-05-22] [WebCite Cache ID 78Z1A5Apj]
25. Cranwell J, Opazo-Breton M, Britton J. Adult and adolescent exposure to tobacco and alcohol content in contemporary YouTube music videos in Great Britain: a population estimate. J Epidemiol Community Health 2016 May;70(5):488-492 [FREE Full text] [doi: 10.1136/jech-2015-206402] [Medline: 26767404]
26. Nasim A, Blank MD, Cobb CO, Berry BM, Kennedy MG, Eissenberg T. How to freak a Black & Mild: a multi-study analysis of YouTube videos illustrating cigar product modification. Health Educ Res 2014 Feb;29(1):41-57 [FREE Full text] [doi: 10.1093/her/cyt102] [Medline: 24162926]
27. Elkin L, Thomson G, Wilson N. Connecting world youth with tobacco brands: YouTube and the internet policy vacuum on Web 2.0. Tob Control 2010 Oct;19(5):361-366. [doi: 10.1136/tc.2010.035949] [Medline: 20739706]
28. Liang Y, Zheng X, Zeng DD, Zhou X, Leischow SJ, Chung W. Exploring how the tobacco industry presents and promotes itself in social media. J Med Internet Res 2015 Jan 21;17(1):e24 [FREE Full text] [doi: 10.2196/jmir.3665] [Medline: 25608524]
29. Carroll MV, Shensa A, Primack BA. A comparison of cigarette- and hookah-related videos on YouTube. Tob Control 2013 Sep;22(5):319-323 [FREE Full text] [doi: 10.1136/tobaccocontrol-2011-050253] [Medline: 22363069]
30. Seidenberg AB, Rodgers EJ, Rees VW, Connolly GN. Youth access, creation, and content of smokeless tobacco. J Adolesc Health 2012 Apr;50(4):334-338. [doi: 10.1016/j.jadohealth.2011.09.003] [Medline: 22443835]
31. Bromberg JE, Augustson EM, Backinger CL. Portrayal of smokeless tobacco in YouTube videos. Nicotine Tob Res 2012 Apr;14(4):455-462 [FREE Full text] [doi: 10.1093/ntr/ntr235] [Medline: 22080585]
32. Richardson A, Vallone DM. YouTube: a promotional vehicle for little cigars and cigarillos? Tob Control 2014 Jan;23(1):21-26. [doi: 10.1136/tobaccocontrol-2012-050562] [Medline: 23047887]
33. Forsyth SR, Malone RE. "I'll be your cigarette—light me up and get on with it": examining smoking imagery on YouTube. Nicotine Tob Res 2010 Aug;12(8):810-816 [FREE Full text] [doi: 10.1093/ntr/ntq101] [Medline: 20634267]
34. Hua M, Yip H, Talbot P. Mining data on usage of electronic nicotine delivery systems (ENDS) from YouTube videos. Tob Control 2013 Mar;22(2):103-106. [doi: 10.1136/tobaccocontrol-2011-050226] [Medline: 22116832]
35. Luo C, Zheng X, Zeng DD, Leischow S. Portrayal of electronic cigarettes on YouTube. BMC Public Health 2014;14:1028 [FREE Full text] [doi: 10.1186/1471-2458-14-1028] [Medline: 25277872]
36. Paek H, Kim S, Hove T, Huh JY. Reduced harm or another gateway to smoking? source, message, and information characteristics of E-cigarette videos on YouTube. J Health Commun 2014;19(5):545-560. [doi: 10.1080/10810730.2013.821560] [Medline: 24117370]
37. Allem J, Escobedo P, Cruz TB, Unger JB. Vape pen product placement in popular music videos. Addict Behav 2017 Nov 3:263-264. [doi: 10.1016/j.addbeh.2017.10.027] [Medline: 29128149]

XSL•FO
RenderX

38. Smith A, Toor S, van Kessel P. Pew Internet. Many turn to YouTube for children's content, news, how-to lessons URL: https://www.pewinternet.org/2018/11/07/many-turn-to-youtube-for-childrens-content-news-how-to-lessons/ [accessed 2019-05-22] [WebCite Cache ID 78Z1gzP32]
39. Freeman B, Chapman S. Is "YouTube" telling or selling you something? Tobacco content on the YouTube video-sharing website. Tob Control 2007 Jun;16(3):207-210 [FREE Full text] [doi: 10.1136/tc.2007.020024] [Medline: 17565142]
40. Freeman B. New media and tobacco control. Tob Control 2012 Mar;21(2):139-144. [doi: 10.1136/tobaccocontrol-2011-050193] [Medline: 22345236]
41. Smith D. Charlie is so 'English'-like: nationality and the branded celebrity person in the age of YouTube. Celebr Stud 2014 Apr 24;5(3):256-274. [doi: 10.1080/19392397.2014.903160]
42. Wang K. Forbes. 2014 Sep 10. The explosive growth of influencer marketing and what it means for you URL: https://www.forbes.com/sites/kylewong/2014/09/10/the-explosive-growth-of-influencer-marketing-and-what-it-means-for-you/ [accessed 2019-05-22] [WebCite Cache ID 78Z1y3xwW]
43. YouTube. YouTube Help URL: https://support.google.com/youtube/answer/2801981 [accessed 2019-05-22] [WebCite Cache ID 78Z2GVpuJ]
44. YouTube. Age-restricted content URL: https://support.google.com/youtube/answer/2802167?hl=en [accessed 2019-05-22] [WebCite Cache ID 78Z2MwQzv]
45. Kim K, Paek H, Lynn J. A content analysis of smoking fetish videos on YouTube: regulatory implications for tobacco control. Health Commun 2010 Mar;25(2):97-106. [doi: 10.1080/10410230903544415] [Medline: 20390676]
46. Jansen BJ, Spink A. How are we searching the World Wide Web? A comparison of nine search engine transaction logs. Inf Process Manag 2006 Jan;42(1):248-263. [doi: 10.1016/j.ipm.2004.10.007]
47. YouTube. 2017. Community Guidelines URL: https://www.youtube.com/yt/about/policies/ [accessed 2019-05-22] [WebCite Cache ID 78Z2Xp2S9]
48. Galstyan E, Galimov A, Sussman S. Commentary: the emergence of pod mods at vape shops. Eval Health Prof 2019 Mar;42(1):118-124. [doi: 10.1177/0163278718812976] [Medline: 30477337]
49. Barrington-Trimis JL, Gibson LA, Halpern-Felsher B, Harrell MB, Kong G, Krishnan-Sarin S, et al. Type of e-cigarette device used among adolescents and young adults: findings from a pooled analysis of 8 studies of 2,166 vapers. Nicotine Tob Res 2018 Jan 05;20(2):271-274. [doi: 10.1093/ntr/ntx069] [Medline: 28371890]
50. Krishnan-Sarin S, Morean M, Kong G, Bold KW, Camenga DR, Cavallo DA, et al. E-Cigarettes and. Pediatrics 2017 Mar;139(3):e20163224 [FREE Full text] [doi: 10.1542/peds.2016-3224] [Medline: 28167512]
51. Chen C, Zhuang Y, Zhu S. E-Cigarette design preference and smoking cessation: a US population study. Am J Prev Med 2016 Sep;51(3):356-363. [doi: 10.1016/j.amepre.2016.02.002] [Medline: 27005984]
52. Talih S, Balhas Z, Salman R, Karaoghlanian N, Shihadeh A. "Direct Dripping": A High-Temperature, High-Formaldehyde Emission Electronic Cigarette Use Method. Nicotine Tob Res 2016 Apr;18(4):453-459 [FREE Full text] [doi: 10.1093/ntr/ntv080] [Medline: 25863521]
53. Kosmider L, Sobczak A, Fik M, Knysak J, Zaciera M, Kurek J, et al. Carbonyl compounds in electronic cigarette vapors: effects of nicotine solvent and battery output voltage. Nicotine Tob Res 2014 Oct;16(10):1319-1326 [FREE Full text] [doi: 10.1093/ntr/ntu078] [Medline: 24832759]
54. Noel JK, Rees VW, Connolly GN. Electronic cigarettes: a new 'tobacco' industry? Tob Control 2011 Jan;20(1):81. [doi: 10.1136/tc.2010.038562] [Medline: 20930060]
55. Berger J, Ward M. Subtle signals of inconspicuous consumption. J Consum Res 2010 Dec 1;37(4):555-569. [doi: 10.1086/655445]
56. Halpern-Felsher BL, Biehl M, Kropp RY, Rubinstein ML. Perceived risks and benefits of smoking: differences among adolescents with different smoking experiences and intentions. Prev Med 2004 Sep;39(3):559-567. [doi: 10.1016/j.ypmed.2004.02.017] [Medline: 15313096]
57. Jamal A, Gentzke A, Hu SS, Cullen KA, Apelberg BJ, Homa DM, et al. Tobacco use among middle and high school students - United States, 2011-2016. MMWR Morb Mortal Wkly Rep 2017 Jun 16;66(23):597-603 [FREE Full text] [doi: 10.15585/mmwr.mm6623a1] [Medline: 28617771]
58. Hu SS, Neff L, Agaku IT, Cox S, Day HR, Holder-Hayes E, et al. Tobacco product use among adults - United States, 2013-2014. MMWR Morb Mortal Wkly Rep 2016;65(27):685-691 [FREE Full text] [doi: 10.15585/mmwr.mm6527a1] [Medline: 27416365]
59. 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]
60. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]

Abbreviations

EDM: electronic dance music

XSL•FO
RenderX

FDA: Food and Drug Administration
NIDA: National Institute on Drug Abuse
NIH: National Institutes of Health
PG: propylene glycol
VG: vegetable glycerin

Edited by G Eysenbach; submitted 08.11.18; peer-reviewed by JP Allem, A Majmundar; comments to author 20.12.18; revised version received 01.03.19; accepted 24.03.19; published 18.06.19.

Please cite as:
Kong G, LaVallee H, Rams A, Ramamurthi D, Krishnan-Sarin S
Promotion of Vape Tricks on YouTube: Content Analysis
J Med Internet Res 2019;21(6):e12709
URL: http://www.jmir.org/2019/6/e12709/
doi:10.2196/12709
PMID:31215510

©Grace Kong, Heather LaVallee, Alissa Rams, Divya Ramamurthi, Suchitra Krishnan-Sarin. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patients' Perceptions of Portal Use Across Care Settings: Qualitative Study

Ann Scheck McAlearney[1,2,3,4], MS, ScD; Cynthia J Sieck[1,2], MPH, PhD; Alice Gaughan[2], MS; Naleef Fareed[2,3], MBA, PhD; Jaclyn Volney[2], BA; Timothy R Huerta[1,2,3,4], MS, PhD

[1]Department of Family Medicine, College of Medicine, The Ohio State University, Columbus, OH, United States

[2]CATALYST (Center for the Advancement of Team Science, Analytics, and Systems Thinking), College of Medicine, The Ohio State University, Columbus, OH, United States

[3]Department of Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, OH, United States

[4]Division of Health Services Management and Policy, College of Public Health, The Ohio State University, Columbus, OH, United States

Corresponding Author:
Ann Scheck McAlearney, MS, ScD
Department of Family Medicine
College of Medicine
The Ohio State University
460 Medical Center Drive, Suite 530
Columbus, OH, 43210
United States
Phone: 1 614 293 8973
Email: Ann.McAlearney@osumc.edu

Abstract

Background: Patient portals are a promising instrument to improve patient-centered care, as they provide patients information and tools that can help them better manage their health. The implementation of portals in both the inpatient and outpatient setting gives health care providers an opportunity to support patients both during hospitalization and after discharge. Thus, there is a need to better understand how inpatient and outpatient portals are used across care contexts.

Objective: This study aimed to examine patients' perceptions of using inpatient and outpatient portals across the care settings, including how they used the portals and the benefits and concerns associated with portal use.

Methods: This study was conducted in a large Midwestern academic medical center consisting of seven hospitals. We interviewed 120 patients who had used an inpatient portal during their hospitalization, at 15 days and 6 months postdischarge, to determine their perspectives of portal use in both hospital and outpatient settings. Interview transcripts were analyzed inductively and deductively by using team coding processes consistent with a grounded theory approach.

Results: Interviews focused on three main areas of portal use: experience with the portal features, perceived benefits, and concerns. Responses at 15 days (n=60) and 6 months (n=60) postdischarge were consistent with respect to perceptions about portal use. Patients identified viewing their health information, managing their schedule, and communicating with providers as notable activities. Convenience, access to information, and better engagement in care were indicated as benefits. Concerns were related to technology issues and privacy/security risks.

Conclusions: Implementation of inpatient portals as a complement to outpatient portals is increasing and can enable patients to better manage aspects of their care. Although care processes vary substantively across settings, the benefits of convenience, improved access to information, and better engagement in care provide opportunities for portal use across care settings to support patient-centered care.

(J Med Internet Res 2019;21(6):e13126) doi:10.2196/13126

KEYWORDS

patient portals; hospitalization; medical informatics; implementation; engagement; ambulatory care

XSL•FO
RenderX

Introduction

The focus on patient-centered health care delivery has made patient portals a promising medium through which patients can gain access to their health information [1,2]. Patients can use the tools and information portals provide to better manage their health, potentially increasing engagement with providers and improving the process of care delivery [3-5].

Inpatient and outpatient portals are usually offered by a health care provider or institution and linked to an electronic health record system. Outpatient portals, typically Web-based, offer patients information focused on outpatient care including access to a health summary, medication listing, immunizations, patient health data entry, appointment tracking, secure messaging, and patient financial account management information. Portals are now increasingly offered in the inpatient setting to hospitalized patients, providing customized information for a patient's inpatient stay including an expected daily care plan, secure messaging with the care team, a notes feature, and educational materials.

Although inpatient and outpatient portals are independent applications, patient information is available through a common electronic health record platform. Together, these two types of portals can provide complete access to health information and providers across care settings. For example, a hospitalized patient could use educational materials, track test results, and make notes during his/her inpatient stay. After discharge, an outpatient portal provides information about the patient's recent hospital stay [6] and can be used to facilitate ongoing management of laboratory and test results as well as continued communication with the patient's care team.

Studies examining portal use primarily focus on a single application: inpatient or outpatient. The evidence on patient use of inpatient and outpatient portals converges on the benefits of portal use, including improved patient-provider interactions, higher awareness of care, and better adherence to care regimens; it also comparably notes portal shortcomings including fears of privacy loss, concerns about results comprehension, and technical challenges with portal use [7]. A recent systematic review of portal use, however, highlighted significantly more evidence regarding outpatient portal use compared with inpatient portals, suggesting the need for more empirical research on the newer inpatient portal technology [8]. Furthermore, there is currently a gap in research that examines the use of patient portals across care settings [9,10]. As such, there remains a clear need for more research to understand both how portal use is evolving in the health care system and how patients perceive portal use across care settings.

This study was conducted in response to the need to understand how patient portals function in different contexts. Our study aims to describe patients' experiences with both inpatient and outpatient portals, including exploring how patients' perceptions compare across settings. This study was approved by The Ohio State University's Institutional Review Board.

Methods

Study Setting

Our study was conducted in a large Midwestern academic medical center consisting of seven hospitals. The center began offering the MyChart outpatient portal in 2011 (Epic Systems, Verona, WI). Patients can access MyChart via the website on a desktop computer or a mobile app on electronic devices (eg, smartphones). In 2013, the academic medical center introduced MyChart Bedside (MCB), a companion portal tailored to the inpatient environment, in a few pilot units before introducing MCB system-wide in 2016. Patients can access MCB on academic medical center–provided tablets during their hospital stay and are able to access their inpatient data via the outpatient portal after discharge.

As shown in Table 1, both MyChart and MCB offer features related to scheduling, personal health information, and health education. In addition, both portals offer slightly different ways to communicate with health care providers. MyChart, the outpatient portal, allows a patient to selectively communicate with an individual provider, either their primary care provider or a specialist; in contrast, messages sent via MCB, the inpatient portal, are sent to the entire care team assigned to that patient. Additional features that are specific to each portal are compared in Table 1.

XSL•FO
RenderX

Table 1. Comparison of features of inpatient and outpatient portals.

Portal feature	Feature description	Inpatient portal function (MyChart Bedside)	Outpatient portal function (MyChart)
Scheduling	View schedule for the day (inpatient) or schedule appointments (outpatient)	Happening Soon	Visits
Health information	View personal health information including biometric data and medications	My Health	My Medical Record
Health education	Read health education materials assigned by health care providers	To Learn	Resources
Secure messaging	Communicate with care team (inpatient) or individual providers (outpatient)	Messages	Messaging
Refill medications	Request refill of prescribed medications	N/A[a]	Request a Refill
Pay bill	Pay inpatient or outpatient bill	N/A	Billing
Set communication preferences	Select frequency of email or text communication	N/A	Preferences
View care team	View names and pictures of all members of inpatient care team assigned to the patient at each shift	Taking Care of Me	N/A
Order meals	Select menu items from the prescribed diet	Dining on Demand	N/A
Request a service or item	Request services such as pastoral care or delivery of a newspaper	I Would Like	N/A
Make notes	Type items the patient wishes to remember	Note to Self	N/A

[a]N/A: not applicable.

Table 2. Participant demographics.

Demographic	Phase 1 (15 days postdischarge)	Phase 2 (6 months postdischarge)
Total participants, n	60	60
Age (years), mean (SD)	49.4 (14.3)	49.4 (13.7)
Male participants, n (%)	18 (30)	15 (25)

Data Collection

We conducted two phases of 15-minute telephone interviews between January 2017 and May 2018, with a total of 120 patients who had been hospitalized at the academic medical center and used the inpatient portal during their recent stay. Phase 1 interviews were completed 15 days postdischarge, and Phase 2 interviews were completed 6 months postdischarge. We chose the timing of these interviews to capture immediate impressions in Phase 1 and longer-term impressions in Phase 2. Furthermore, the timing of these interviews provided a window during which perspectives about both inpatient portal use (during hospitalization) and outpatient portal use (postdischarge) could be evaluated. Interview phases involved samples of different patients (Table 2), randomly selected from among all discharged patients who had consented to study participation. Patients were recruited for interviews based on their discharge dates within the appropriate time windows. We did not select our sample based on the level of portal use, as we wanted to understand the perspectives of all types of portal users. In addition, we intended to examine a range of patients' perceptions about the portals, not limited by demographic characteristics or conditions.

Study investigators conducted interviews using semistructured guides designed to explore patient experiences with both the inpatient and outpatient portals. Patient interview guides are provided in Multimedia Appendix 1. These guides were developed by the research team and pilot tested on a sample of patient volunteers. The 15-day guide focused on the use of and experience with the inpatient portal during the patient's recent hospital stay; the 6-month guide examined how the inpatient portal experience influenced outpatient portal use and asked about long-term use of the outpatient portal. All interviews were audio-recorded, transcribed verbatim, and de-identified for analysis.

Data Analysis

Interview transcripts were analyzed both inductively and deductively, in accordance with rigorous qualitative methods [11,12]. First, a preliminary coding dictionary was developed based on questions asked in the interview guide. Next, a sample of four interview transcripts were coded by four members of the research team (AM, CS, AG, and TW) to refine the coding dictionary and explore the emergence of new codes in the data. Using the refined coding dictionary, all interview transcripts were coded by two members of the research team who held frequent meetings to ensure consistency of coding and agreement about the definitions of new codes as they emerged during the coding process, consistent with our grounded theory approach [13]. Development and refinement of the coding dictionary and all aspects of the coding and analysis were led

XSL•FO
RenderX

by an experienced qualitative researcher (AM). ATLAS.ti (version 6.0; ATLAS.ti Scientific Software Development GmbH, Berlin, Germany) qualitative data analysis software was used to support this analysis.

Results

Overview

Three main topics were discussed by patients when they were asked about their experience with patient portals: experience with the portals, perceived benefits of portal use, and concerns about portals. Patients' responses to interview questions at 15 days and 6 months postdischarge did not vary appreciably from each other; as a result, the findings we present do not distinguish between interview time frames.

Patients' Experiences with Portal Use

When describing their use of portals across care settings, patients primarily discussed three main activities: viewing health information, managing schedules, and communicating with providers. Patients' uses of these features are discussed below with additional exemplary quotes presented in Textbox 1.

Textbox 1. Patients' experiences with portal use (portal features used and examples of use).

Viewing health information

Examples of inpatient portal use:

- I just used it to look at what my medicines were and what kind of, and then what the medicines were all about, what was in them and everything like that. [Patient #610]
- I got to look at my charts and see my bloodwork and see how things were going. [Patient #627]

Examples of outpatient portal use:

- To find out my medical condition, the symptoms…any details about a medical condition, you can pretty much trace it back and go through that system, and find out things. [Patient #613]
- I can now see test results and also just whatever medical information that's helpful. [Patient #610]

Managing schedules

Examples of inpatient portal use:

- I could kind of know what the plan was gonna be the days that I was in the hospital. [Patient #623]
- The ability to stay on track with what your schedule is and know what your treatment plan is. [Patient #620]

Examples of outpatient portal use:

- I'll look up my appointments sometimes if I have a scheduling conflict then I utilize that tab for a more direct access to my providers. [Patient #619]
- I use it to make appointments and change appointments...[Patient #624]

Communicating with providers

Examples of inpatient portal use:

- What I liked about it was that I was able to correspond with the doctors, nurses. [Patient #614]
- I messaged my doctors to see what the…cause like when you look at your bloodwork it's got this scale and it's got numbers and I didn't know what they were. So I, it let me hit a button and ask them questions: "What does this mean?" and "Am I okay? Am I going to live?" [Patient #627]

Examples of outpatient portal use:

- I can e-mail the doctors and they, you know, e-mail me back. I think it's a very good tool. [Patient #608]
- Well, it's communication with my doctor, my regular doctor. He'll put notes in there and I can put notes back to him. [Patient #626]

Viewing Health Information

Views of laboratory and test results as well as features to learn more about what these results meant were important portal features used by patients. In the inpatient setting, patients could view results as soon as they became available and track multiple tests throughout the day. For example, one patient reflected,

> *I used it daily, probably a few times a day. When I have a new test, I just check when it's there, or if the results are back or not.* [Patient #104]

In the outpatient environment, patients could also view results as soon as they were available, but their portal use also allowed them to monitor changes over time:

> *I use it to look up new lab results when they're available, and also sometimes to look up older ones to kind of look at the trends.* [Patient #624]

XSL•FO
RenderX

Managing Schedules

Both in the hospital and outpatient settings, patients reported using the schedule feature. Inpatients appreciated knowing when tests and medication administrations were scheduled:

> *I liked the schedule that was on there of what the various times of when I was going to receive my medicine and everything. That way I knew what was coming.* [Patient #610]

In the outpatient setting, patients noted frequently using this scheduling feature, "to keep track of appointments" (Patient #228).

Communicating With Providers

The ability to communicate with providers was a feature used in both portals. One patient described their use of the secure messaging feature while hospitalized:

> *If you didn't understand something from the doctor, you could message him and ask him.* [Patient #625]

In the inpatient environment, however, not all patients reported the need for another way to communicate with their providers, given the frequency of in-person interactions with their care team:

> *They [care team members] came in every day and told me what the test results were, so I didn't spend a lot of time on that.* [Patient #112]

In contrast, in the outpatient environment, the majority of patients used the portal to communicate with providers:

> *Well, since I've gotten out of the hospital, just to stay in contact with my doctor, or doctor's visits that I might not even have to see because I was able to communicate with him and find things out without going to the doctor's office.* [Patient #626]

Perceived Benefits of Portal Use

Patients in our study noted benefits from their use of portals in both the inpatient and outpatient settings, and these centered around three themes: convenience, improved access to information, and better engagement in care. Below, we describe the perceived benefits in these areas and present additional supporting quotations in Textbox 2.

XSL•FO
RenderX

Textbox 2. Perceived benefits of portal use (with examples of use).

Convenience

Examples of inpatient portal use:

- I'd say it was nice because it made it easier for me to, pretty much just whenever I wanted to, or whenever it was convenient for me. [Patient #202]
- That you could see your lab results and medications and everything. I liked it. I mean it was easy to use. [Patient #625]

Examples of outpatient portal use:

- Well MyChart to me is efficient. It's what I need to get through my day and I can do it at work, anywhere I'm at any time of day. So it's extremely beneficial. [Patient #626]
- You do not have to be put on hold for MyChart. It's all just right there and there is no other step or hoop to jump through to get the answers that you need for that. [Patient #607]

Improved access to information

Examples of inpatient portal use:

- The knowledge of what was actually going on and keeping up with my medications as they were changing and stuff. All that. That's what I really liked about it was the ability to help me keep up. [Patient #617]
- One thing that I just remembered that I really liked was like it would tell me like when I might be getting discharged and how many days I've been in the hospital as well because sometimes you tend to lose track. [Patient #227]

Examples of outpatient portal use:

- I liked it so well is because my health was quickly changing and I really didn't understand all of why it was changing so quickly. And it helped me to understand those changes. [Patient #617]
- I needed to remember appointments because I was having appointments here and there and all over the hospital and all over campus. I was all over campus and I needed to remember what time, what day where I was supposed to be, and then what the results were from each appointment that I had at that time. [Patient #618]

Better engagement in care

Examples of inpatient portal use:

- I was able to read, read my test results, usually before they [the care team] came in, and I was able to figure out questions I wanted to ask them before they got there. [Patient #109]
- I feel like I'm an active participant. [Patient #229]

Examples of outpatient portal use:

- Every time I open up my email, I always pay attention and make sure that I don't miss the emails saying that yeah when you message or you have something coming from MyChart. And then I immediately go and log in and see what areas and what has been added like new test results and new appointments. So that is always done to keep up to date and it is also extremely helpful. [Patient #122]
- It's more just being one-on-one with the doctors, you know. If I need a refill I can e-mail the doctor or nurse or whatever and get it done or seeing results from tests, seeing my appointments. It's a big help. [Patient #608]

Convenience

Patients noted many ways in which the portals in different care settings were convenient tools they could use on their own time. In the hospital, patients appreciated the ability to order their own meals:

> *Ordering food from [Dining on Demand]…that was all extremely helpful because…you have time to look at everything that's there, and it allows you by far a much more varied menu from meal to meal, especially if you stay more than just a couple of days.* [Patient #122]

Another feature patients found convenient was the ability to review their own health information without waiting for a provider:

> *I could just log into it and look and see the results versus having to wait for the doctor or the nurse to come in and tell me the results.* [Patient #107]

The convenience of the outpatient portal was similarly valued, especially in comparison to not having access to a portal:

> *I find it far more convenient. Like when I'm waiting for lab results, to know that when that result is available, I can readily see it on there. I don't have to wait potentially days for the doctor's office to have time to go through all their results and contact me, and then play phone tag with each other.* [Patient #624]

Another patient similarly explained,

XSL•FO
RenderX

> *I think having one central place to go to, to look up things like lab results and past appointments and surgery stuff is very helpful to have all in one place.* [Patient #607]

Improved Access to Information

Patients noted that portals provided improved access to information about their current health care plans and medical histories. For example, one patient described the benefits of having ready access to information while in the hospital:

> *Well it keeps you more in tune to your health. I mean it tells you, you got meds coming. It tells you why you're taking the meds. There's advice that's good for everybody. Cause doctors fail to realize that we don't understand all that medical terminology. Once they leave, I click on that little thing and it tells you what it's for and…it tells you what it is and what the symptoms of it are and what they can do to fix it.* [Patient #604]

Similarly, in the outpatient setting, patients described using the outpatient portal to review discussions with their doctor:

> *The communication is very beneficial. I can also double check if I can't remember what the doctor had explained with test results and things like that. I can always…look back at those.* [Patient #621]

Better Engagement in Care

Patients also reported that using the portals helped them fully engage with their care. For example, one patient told us that during their hospital stay, "It [MCB] gave me the information to ask the right questions" (Patient #209). Another patient specifically described the benefit of using the inpatient portal in contrast to not having it available:

> *You hit your call light. The nurse or the nurse assistant that comes in and they're like, "What do you need?" But you don't have a direct link to ask your doctor [like you do with the portal]…You have to wait 24 hours or so for the doctor to come down and see you. But that doctor is only going to be in there three or four minutes, not even four minutes, talking with you.* [Patient #211]

In the outpatient setting, patients similarly described how portals helped them engage in and take control of their health:

> *If it's a person in charge of their own healthcare, and it's not all responsibility of the medical facility anymore. But now it's my responsibility because I have resources that I can go to that I can use.* [Patient #613]

Another patient explicitly noted the benefit of portal access across care settings:

> *I knew that once I was released from the hospital, I could go back, and if I wanted to, consult my doctors or send them a message about what my health was doing. So in that way it was nice.* [Patient #156]

Concerns About Portal Use

There were two main categories of concerns reported by patient participants: technology-related issues and privacy and security issues.

Technology-Related Issues

Patients raised concerns about technology-related issues with both the MCB portal and the tablets. Several patients noted issues with the battery life of the tablets on which the portal was hosted. One patient reflected on the need for internet access outside the hospital to facilitate use of the outpatient version of the portal:

> *...if you're a person that, you know, moves about a lot and maybe doesn't have a…well you have to have internet for one thing. That's one of the things, reasons why I didn't actually use it is when I'm at my place, I don't have internet. So unless I go to a library or I'm at my parents', I don't have any way of getting online.* [Patient #602]

Interestingly, patients occasionally suggested ways to improve the available features on the inpatient portal application. For example, one patient noted that the portal could be improved by providing patients the ability to send a secure message to a specified individual:

> *I wanted a specific doctor, or something to ask them a question, or a specific nurse or something. But it would always send it as a group message, and they said that they didn't know who would respond to me.* [Patient #130]

Privacy and Security Issues

Patients also generally noted security and privacy issues related to portals:

> *Just like anything else, you know, you're on the computer, people get in and get your business. So that's the only thing I was worried about.* [Patient #116]

In the hospital setting, patients expressed concerns about using a hospital-provided tablet that would have to be returned and then used by another patient:

> *The one concern I had was that it was going to be wiped. So I took that…when I was finished with it, since it had my personal Kindle stuff on there. The one thing that could've been done, that sort of would've made me more confident, and I was confident enough to give it to them, was that actually had a button on the app…on the device itself, that I could erase it.* [Patient #110]

Discussion

Principal Findings

Portals have recently been incorporated across most aspects of patient care [14,15]. Although studies of outpatient portals have demonstrated that use can improve acceptance of preventive screening and adherence to disease management plans [16-19], the introduction of an inpatient portal can address the particular

XSL•FO
RenderX

needs of a patient during hospitalization, such as the ability to communicate with the care team and know when to expect medications or treatments.

Patients in our study reported various ways in which they used portals across care settings as they became and remained involved in their care. For example, study participants noted ways in which increased access to information helped them better understand their care plan, such as through the use of scheduling features of both the inpatient and outpatient portals. While hospitalized, the inpatient portal informed patients when to expect medications or tests; once discharged, the outpatient portal allowed them to track appointments and medication refills. In both of these situations, participants noted that a better understanding of the flow of their care made them feel more engaged in the care process. The same appreciation for increased access to information through the portal was noted by patients who followed their laboratory results; patients reported valuing the ability to monitor results, as the results changed during their hospital stay, as well as the ability to track results over time in the outpatient portal. Together, this ecosystem of portals appears to help patients better understand their health trajectory, from moments of crisis (ie, hospitalization) to patterns of management (ie, outpatient care).

Our findings can be viewed using the Umar and Mundy model of patient empowerment that views health information technology as an underlying supportive mechanism for empowerment. Key elements of this model include patient access to information, knowledge development, patient participation as a partner, and functions that facilitate a sense of self-efficacy and align with our findings [20]. In our study, across settings, increased access to information appears to improve the experience of care as well as help patients to engage more fully in the care process, consistent with Umar and Mundy's model that positioned patient portals as health information technology that supports empowerment. For instance, in the inpatient setting, lack of information has been identified as a concern that leads to increased anxiety for patients [8,21]. Inpatients in our study specifically noted the benefits of improved access to information, explaining that having access to the portal allowed them to keep track of the care process, thereby reflecting greater empowerment. In the outpatient setting, access to information is also critical for facilitating communication both in-person as well as between visits, with recent research showing that patient-provider in-person communication is enhanced when patients have viewed the results prior to their visit [22]. Communication with providers, which helps the patient become more of a partner in their care, is also an element identified in Umar and Mundy's model that supports patient empowerment.

Our study is the first to explore patient perspectives of how inpatient and outpatient portals are used across care settings [2]. These findings provide preliminary evidence about the ways in which portal availability influences patients' care experiences, noting convenience and access to information as benefits of both the inpatient and outpatient portals. As such, our results can serve as a foundation to an evidence-based approach that informs the development of portals that promote health, as recommended by Osborn and colleagues [23]. Further, patients described how portal use helped them stay engaged with and manage their health care, suggesting that portal access may have an important impact on factors such as patient self-efficacy [24,25], another important element of patient empowerment identified by Umar and Mundy [20], and the overall care experience [24,25] as patients attempt to better manage their health across care settings.

These study results are consistent with prior research that has highlighted barriers to portal use such as the need for technical support, training, and improved usability [23,26]; thus, addressing these barriers will be an important consideration as portal use extends across the health care continuum. Moreover, our study participants appreciated both the availability of the secure messaging feature and the speed with which providers could respond to their questions, consistent with the results of prior work that has found this feature useful [15,27,28]. At the same time, our study also highlights a continued need to address access to health information technology across care settings. Patients valued the portals, but some patients noted that, due to lack of internet access in the outpatient setting, their most consistent access to the portal was during hospitalization. Studies have found that approximately 70% of the US population has internet access [29], but such reports also note that access is strongly correlated with sociodemographic factors. Research has documented that disparities in internet access and use contribute to poorer health outcomes when comparing outcomes of individuals with consistent access [30-33], thus suggesting the importance of promoting health information technology access to improve health and health care.

Stakeholders in the health care system recognize the value of providing an integrated portal experience that spans care settings [14]. However, there are currently few theories or studies that effectively describe patients' uses of technologies, and fewer still that can take into consideration their distinct needs across the health care continuum [26,34]. We found that having access to an inpatient portal integrated with an outpatient portal helped patients with the care process by providing different applications across care settings, from preventive care to diagnosis and treatment to disease management. Continuing to improve our understanding of the impact and influence of this new technology will be critical as we attempt to advance both the quality and patient-centeredness of care across care settings.

We noted several potential limitations to our study. First, our study examined patient perceptions related to the use of portals in a single health care system. Although the features available are common among most portals offered at other health care systems, policies regarding implementation and use may result in different experiences at different systems, and different portals in use may have features we could not study. Second, it is possible that patients' recalls of their use of the portals were influenced by the time that had elapsed since their hospitalization. However, given the common ideas expressed at both 15 days and 6 months postdischarge, we believe the impact of time is likely minimal; instead, the stability of these findings across time may be a strength of our study. Third, patients who agreed to participate in our interviews might represent the most engaged group of patients. Although their use and perceptions may vary from those of less engaged

XSL•FO
RenderX

patients, their insights are useful to inform future research and practical efforts to introduce and promote portal use. Finally, this qualitative study did not link patients' comments to information about their hospital stay (eg, length of stay or location of treatment) or to their prior portal experience (eg, outpatient portal use prior to hospitalization). Further work that considers these dimensions of care and experience is needed to draw conclusions about the types of patients or clinical situations that can best benefit from different aspects of patient portal use.

Future studies of portal use should consider topics such as portal use patterns, including the relationships between concerns about portal use and subsequent use. There may also be idiosyncrasies of patient portals (eg, the timeliness of information availability in the outpatient portal after an inpatient discharge) that impact portal use across care settings, suggesting opportunities for continued study in this area. Further, examining the use of specific portal features and taking into account patient demographics as well as patient stage along the care continuum would be informative. As health care systems and researchers gain experience with portals that can span the continuum of care, it will be important to take advantage of opportunities to explore portal use in greater detail using both qualitative and quantitative methods.

Conclusions

Development and implementation of inpatient portals as a companion to outpatient portals represent a nascent approach to portal use across the continuum of care. Although the processes of care are substantively different between inpatient and outpatient settings, among those who use inpatient portals, there is a clear preference for this technology. Common functions across the portal tools create familiarity and, by extension, increase engagement. Further, while these tools enable communication through messaging features, they are more useful because of the important information they provide patients to help them in their care journey. Convenience, improved access to information, and better engagement in care suggest specific aspects of portal use that can support patient engagement across care settings.

Acknowledgments

The authors wish to thank Lindsey Sova, Danijela Cvijetinovic, Toby Weinert, Allison Silverman, Ayanna Scott, and Karen Alexander, all affiliated with the authors' organization, for their assistance with this study. This research was reviewed and approved by The Ohio State University's Institutional Review Board. This work was supported by grants from the Agency for Healthcare Research and Quality (R01HS024091, R21HS024767, and P30HS024379). This funding source had no role in the preparation, review, or approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Semistructured patient interview guides.

[PDF File (Adobe PDF File), 129KB - jmir_v21i6e13126_app1.pdf]

References

1. Giardina TD, Baldwin J, Nystrom DT, Sittig DF, Singh H. Patient perceptions of receiving test results via online portals: a mixed-methods study. J Am Med Inform Assoc 2018 Apr 01;25(4):440-446. [doi: 10.1093/jamia/ocx140] [Medline: 29240899]
2. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
3. Sieck CJ, Hefner JL, McAlearney AS. Finding the patient voice in research activities: perspectives of patient advisors and researchers. J Participat Med 2017 Apr 01;9:e6 [FREE Full text]
4. Peacock S, Reddy A, Leveille SG, Walker J, Payne TH, Oster NV, et al. Patient portals and personal health information online: perception, access, and use by US adults. J Am Med Inform Assoc 2017 Apr 01;24(e1):e173-e177. [doi: 10.1093/jamia/ocw095] [Medline: 27413120]
5. Ahern DK, Woods SS, Lightowler MC, Finley SW, Houston TK. Promise of and potential for patient-facing technologies to enable meaningful use. Am J Prev Med 2011 May;40(5 Suppl 2):S162-S172. [doi: 10.1016/j.amepre.2011.01.005] [Medline: 21521591]
6. McAlearney AS, Sieck CJ, Hefner JL, Aldrich AM, Walker DM, Rizer MK, et al. High touch and high tech (HT2) proposal: transforming patient engagement throughout the continuum of care by engaging patients with portal technology at the bedside. JMIR Res Protoc 2016 Nov 29;5(4):e221 [FREE Full text] [doi: 10.2196/resprot.6355] [Medline: 27899338]
7. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012 Nov 26;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]

XSL•FO
RenderX

8. Kelly MM, Coller RJ, Hoonakker PL. Inpatient portals for hospitalized patients and caregivers: a systematic review. J Hosp Med 2018 Jun 01;13(6):405-412. [doi: 10.12788/jhm.2894] [Medline: 29261819]
9. Dalal AK, Bates DW, Collins S. Opportunities and challenges for improving the patient experience in the acute and postacute care setting using patient portals: the patient's perspective. J Hosp Med 2017 Dec;12(12):1012-1016. [doi: 10.12788/jhm.2860] [Medline: 29073310]
10. Reti SR, Feldman HJ, Ross SE, Safran C. Improving personal health records for patient-centered care. J Am Med Inform Assoc 2010;17(2):192-195 [FREE Full text] [doi: 10.1136/jamia.2009.000927] [Medline: 20190063]
11. Crabtree BF, Miller WL. Using codes and code manuals: a template organizing style of interpretation. In: Doing qualitative research. Thousand Oaks, CA: SAGE Publications; 1999:163-177.
12. Constas MA. Qualitative analysis as a public event: the documentation of category development procedures. Am Educ Res J 1992 Jun 01;29(2):253-266. [doi: 10.3102/00028312029002253]
13. Glaser BG, Strauss AL, Strutzel E. The discovery of grounded theory; strategies for qualitative research. Nursing Research 1968;17(4):364. [doi: 10.1097/00006199-196807000-00014]
14. Collins SA, Rozenblum R, Leung WY, Morrison CR, Stade DL, McNally K, et al. Acute care patient portals: a qualitative study of stakeholder perspectives on current practices. J Am Med Inform Assoc 2017 Apr 01;24(e1):e9-17. [doi: 10.1093/jamia/ocw081] [Medline: 27357830]
15. Sieck CJ, Hefner JL, Schnierle J, Florian H, Agarwal A, Rundell K, et al. The rules of engagement: perspectives on secure messaging from experienced ambulatory patient portal users. JMIR Med Inform 2017 Jul 04;5(3):e13 [FREE Full text] [doi: 10.2196/medinform.7516] [Medline: 28676467]
16. Wald JS, Middleton B, Bloom A, Walmsley D, Gleason M, Nelson E, et al. A patient-controlled journal for an electronic medical record: issues and challenges. Stud Health Technol Inform 2004;107(Pt 2):1166-1170. [Medline: 15360996]
17. Wilson ML, Murphy LS, Newhouse RP. Patients' access to their health information: a meaningful-use mandate. J Nurs Adm 2012 Nov;42(11):493-496. [doi: 10.1097/NNA.0b013e3182714538] [Medline: 23099997]
18. Halamka JD, Mandl KD, Tang PC. Early experiences with personal health records. J Am Med Inform Assoc 2008;15(1):1-7 [FREE Full text] [doi: 10.1197/jamia.M2562] [Medline: 17947615]
19. Coughlin SS, Prochaska JJ, Williams LB, Besenyi GM, Heboyan V, Goggans DS, et al. Patient web portals, disease management, and primary prevention. Risk Manag Healthc Policy 2017 Apr 07;10:33-40 [FREE Full text] [doi: 10.2147/RMHP.S130431] [Medline: 28435342]
20. Umar A, Mundy D. Re-thinking models of patient empowerment. Stud Health Technol Inform 2015;209:175-181. [Medline: 25980722]
21. Yoo S, Lee KH, Baek H, Ryu B, Chung E, Kim K, et al. Development and user research of a smart bedside station system toward patient-centered healthcare system. J Med Syst 2015 Sep;39(9):86. [doi: 10.1007/s10916-015-0273-8] [Medline: 26208595]
22. White A, Danis M. Enhancing patient-centered communication and collaboration by using the electronic health record in the examination room. JAMA 2013 Jun 12;309(22):2327-2328 [FREE Full text] [doi: 10.1001/jama.2013.6030] [Medline: 23757080]
23. Osborn CY, Mayberry LS, Wallston KA, Johnson KB, Elasy TA. Understanding patient portal use: implications for medication management. J Med Internet Res 2013 Jul 03;15(7):e133 [FREE Full text] [doi: 10.2196/jmir.2589] [Medline: 23823974]
24. Lorig KR, Sobel DS, Ritter PL, Laurent D, Hobbs M. Effect of a self-management program on patients with chronic disease. Eff Clin Pract 2001;4(6):256-262. [Medline: 11769298]
25. Sorondo B, Allen A, Fathima S, Bayleran J, Sabbagh I. Patient portal as a tool for enhancing patient experience and improving quality of care in primary care practices. EGEMS (Wash DC) 2017 Jan 26;4(1):1262 [FREE Full text] [doi: 10.13063/2327-9214.1262] [Medline: 28203611]
26. Winkelman WJ, Leonard KJ, Rossos PG. Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. J Am Med Inform Assoc 2005;12(3):306-314 [FREE Full text] [doi: 10.1197/jamia.M1712] [Medline: 15684128]
27. Nazi KM. The personal health record paradox: health care professionals' perspectives and the information ecology of PHR systems in organizational and clinical settings. 2012 Presented at: Medicine 2.0 Conference; September 15-16, 2012; Boston, MA.
28. Houston TK, Sands DZ, Jenckes MW, Ford DE. Experiences of patients who were early adopters of electronic communication with their physician: satisfaction, benefits, and concerns. Am J Manag Care 2004 Sep;10(9):601-608 [FREE Full text] [Medline: 15515992]
29. Cohn D. Pew Research Center. 2014. Census: computer ownership, internet connection varies widely across US URL: https://www.pewresearch.org/fact-tank/2014/09/19/census-computer-ownership-internet-connection-varies-widely-across-u-s/ [accessed 2019-05-13] [WebCite Cache ID 78LqlKV5I]
30. Estacio EV, Whittle R, Protheroe J. The digital divide: examining socio-demographic factors associated with health literacy, access and use of internet to seek health information. J Health Psychol 2017 Feb 01:1359105317695429. [doi: 10.1177/1359105317695429] [Medline: 28810415]

XSL•FO
RenderX

31. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: internet use patterns, ehealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013 May 02;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
32. Aydın G, Kaya N, Turan N. The role of health literacy in access to online health information. Procedia - Social and Behavioral Sciences 2015 Jul 03;195:1683-1687. [doi: 10.1016/j.sbspro.2015.06.252]
33. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. J Gen Intern Med 2011 Oct;26(10):1117-1123 [FREE Full text] [doi: 10.1007/s11606-011-1749-y] [Medline: 21647748]
34. Or CK, Karsh BT. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]

Abbreviations

MCB: MyChart Bedside

Edited by G Eysenbach; submitted 13.12.18; peer-reviewed by F Bourgeois, D Alhuwail, T Irizarry, R Dendere; comments to author 14.02.19; revised version received 01.03.19; accepted 24.03.19; published 06.06.19.

Please cite as:
McAlearney AS, Sieck CJ, Gaughan A, Fareed N, Volney J, Huerta TR
Patients' Perceptions of Portal Use Across Care Settings: Qualitative Study
J Med Internet Res 2019;21(6):e13126
URL: https://www.jmir.org/2019/6/e13126/
doi:10.2196/13126
PMID:31172960

©Ann Scheck McAlearney, Cynthia J Sieck, Alice Gaughan, Naleef Fareed, Jaclyn Volney, Timothy R Huerta. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Use of an Electronic Health Record Patient Portal to Access Diagnostic Test Results by Emergency Patients at an Academic Medical Center: Retrospective Study

Brody Foster[1], BS; Matthew David Krasowski[1], MD, PhD

Department of Pathology, University of Iowa Hospitals and Clinics, Iowa City, IA, United States

Corresponding Author:
Matthew David Krasowski, MD, PhD
Department of Pathology
University of Iowa Hospitals and Clinics
200 Hawkins Drive
C-671 GH
Iowa City, IA, 52242
United States
Phone: 1 3193849608
Fax: 1 3193848051
Email: mkrasows@healthcare.uiowa.edu

Abstract

Background: Electronic health record (EHR) patient portals provide a means by which patients can access their health information, including diagnostic test results. Little is known about portal usage by emergency department (ED) patients.

Objective: The study aimed to assess patient portal utilization by ED patients at an academic medical center using account activation rates along with the rates of access of diagnostic test results (laboratory results and radiology reports), analyzing the impact of age, gender, and self-reported patient race.

Methods: This institutional review board–approved retrospective study was performed at a 60,000-visits-per-year university-based ED. We utilized EHR data reporting tools to examine EHR portal activation and utilization for all patients who had at least one ED encounter with one or more diagnostic tests performed between October 1, 2016, and October 1, 2017. The total dataset for laboratory testing included 208,635 laboratory tests on 25,361 unique patients, of which 9482 (37.39%) had active portal accounts. The total dataset for radiologic imaging included 23,504 radiology studies on 14,455 unique patients, of which 5439 (37.63%) had an active portal account.

Results: Overall, 8.90% (18,573/208,635) of laboratory tests and 8.97% (2019/22,504) of radiology reports ordered in the ED were viewed in the patient portal. The highest rates of viewing of laboratory and radiology results were seen for those who were female, were aged 0 to 11 years (parent or guardian viewing by proxy) and 18 to 60 years, and self-reported their race as Caucasian or Asian. The lowest rates were for those who were teenagers, aged older than 81 years, African American/black, and Hispanic/Latino. Infectious disease, urinalysis, and pregnancy testing constituted the highest number of laboratory tests viewed. Magnetic resonance imaging reports were viewed at higher rates than computed tomography or x-ray studies (P<.001). Approximately half of all the diagnostic test results accessed by patients were reviewed within 72 hours of availability in the patient portal (laboratory results: 9904/18,573, 53.32% and radiology reports: 971/2019, 48.1%). On the other extreme, 19.9% (3701/18,573) of laboratory results and 31.6% (639/2019) of radiology reports were viewed more than 2 weeks after availability in the portal.

Conclusions: The data highlight the relatively low use of a patient portal by ED patients and existing disparities between patient groups. There can be wide lag time (months) between result/report availability and access by patients. Opportunities for improvement exist for both activation and more robust utilization of patient portals by ED patients.

(J Med Internet Res 2019;21(6):e13791) doi:10.2196/13791

KEYWORDS

consumer health informatics; diagnostic imaging; electronic health records; medical informatics; minority groups; patient portals; proxy; user-computer interface

XSL•FO
RenderX

Introduction

Context

Electronic health records (EHRs) are common among many industrialized countries and provide a way for patient health information to be stored and accessed in an efficient and secure fashion [1-4]. EHRs can also provide a tool for patients to access their own health information. Web-based EHR portals (also known as *patient portals*) have become increasingly popular over the past decade [5-11] and have been promoted by legislative and regulatory initiatives such as the Health Information Technology for Economic and Clinical Health (HITECH) act in the United States [12-14]. HITECH aims to improve the overall quality of health care in the United States by encouraging patient involvement in medical decisions. HITECH has also included financial incentives and penalties that have led to EHR implementation in over 86% of US physician practices [1]. Health care institutions have found EHR patient portals to be a popular feature for patients [5,9,10,15-17]. Patient portals can theoretically reduce time-consuming tasks for the health care team such as phone calls (especially if patients are not readily accessible) and mailing hard copies of letters.

Background

Previous studies have shown that patients have a positive perception of using patient portals to become more fully engaged in their own medical decisions [5,10,15-18]. Patient portal use has also been linked to improvements in health care coordination, decreased costs, and better communication between patients and health care providers [19-22]. Appointment scheduling, medication refills, pathology results, radiology reports, and direct messaging with the health care team are some of the most popularly accessed features in patient portals [5,8,9,16,23-25]. Multiple studies in the United States and Europe have shown that the highest users of patient portals tend to be Caucasian, female, and adults aged 25 to 50 years [8,9,16,23,26-30]. However, there are factors that can impact acceptance and use of patient portals. These include language fluency, health care literacy, comfort and familiarity with using the internet, access to broadband internet access, support and assistance from family and friends, and encouragement by the health care team [6,30-35]. The published literature on patient portals has mostly focused on the use of these portals in the outpatient setting, where the full functionality of the portals can be utilized (eg, messaging with providers, prescription refills, scheduling appointments, and viewing provider reports and diagnostic test results). In addition to outpatient applications, some recent literature highlights the use of patient portals in the inpatient setting (eg, sending questions to the health care team and ordering hospital meals) [34,36-39].

In contrast, there have been few studies of patient portal usage by emergency department (ED) patients [40,41]. A previous survey indicated ED physicians' concerns with the release of diagnostic test results to patients while acknowledging some benefits [36]. The population of patients seen in the ED presents some challenges for promoting patient portal usage as these patients may not seek regular medical care at the institution or health network affiliated with the ED. Popular patient portal features such as messaging with the clinical team, pharmacy refills, and appointment scheduling may be of little or no benefit to ED patients. On the contrary, the use of a patient portal affords patients easy access to diagnostic test results and empowers patients to transfer information to other health care facilities. Perhaps, most importantly, portals can allow ED patients to review diagnostic test results and other EHR information (eg, discharge summaries) in their own time; this may be especially advantageous for ED visits for medical issues that can impact cognition and understanding during the acute event.

We have previously analyzed general patterns of patient access of diagnostic test results through the patient portal (Epic MyChart) used at our institution, a Midwestern academic medical center (University of Iowa Hospitals and Clinics; UIHC) [41,42]. One previous study analyzed broad patterns of diagnostic test release and patient access, in particular, focusing on the mechanism and timing of results transmittal to the patient portal [41]. This previous study provides details on the release categories for diagnostic tests and also describes the challenges and potential safety issues encountered at UIHC with patient access of their own test results/reports. A later study focused in detail on outpatient use of the patient portal and identified disparities in portal activation and usage by underrepresented minorities [42].

Objectives

In this study, we analyzed the activation and use of a patient portal by patients at our 60,000-visits-per-year, university-based ED that serves as a regional level 1 trauma center. We focused on patient viewing of laboratory results and radiology reports (collectively referred to as *diagnostic tests*) ordered within the ED, as these are patient portal functions frequently used by ED patients at our institution. For viewing of the diagnostic test results, an analysis was performed for both the total cohort (including those who have never activated a patient portal account) and the more limited subset with active patient portal accounts. Analysis of the 2 populations helps address the separate impact of 2 broad barriers in patient portal usage: getting patients to activate accounts and, once activated, to utilize portal functionality such as viewing diagnostic test results.

Some popular functions of the patient portal in the UIHC outpatient setting are not applicable to the ED patient population. In particular, scheduling of ED visits is not available in the UIHC portal. Prescription refills via the patient portal are also not applicable as ED prescriptions at UIHC are almost always written without refills. In addition, messaging of the health care team via the portal, although possible for ED patients, has not been offered as a preferred means of ED follow-up communications compared with telephone calls. In the period of retrospective analysis, ED patient messaging for follow-up questions was of negligible volume (<15/month). Our hypothesis was that utilization of the patient portal for diagnostic test result viewing in the ED would be lower than previously published rates in our outpatient population but would show similar variation based on age, gender, and self-reported race.

XSL•FO
RenderX

Methods

Study Setting

This study was a retrospective analysis of EHR patient portal records at the UIHC ED between October 1, 2016, and October 1, 2017. The institution has 811 inpatient beds at the main campus and outpatient clinics throughout the local geographic region. UIHC is a tertiary/quaternary care facility that serves as a regional center for specialty care.

Patient Portal

The institution uses Epic (Epic, Inc) as the EHR and adopted its associated patient portal (MyChart) in 2010 [41]. To set up a patient portal account, patients receive information in their after-visit summary documents that include a MyChart activation code and instructions for activating and using their portal account. Patients may also request an activation code online. Parents and legal guardians can activate and manage portal accounts of children aged 11 years and younger by proxy with full functionality. Proxy functionality is limited for children aged 12 to 17 years (eg, parents and guardians cannot access diagnostic test results) and ends when children turn 18 years. With documentation, legal guardians of adult dependent patients may obtain proxy access.

The MyChart graphical user interface has evolved over time, but here, we summarize some of the basic elements of the one in use at UIHC during the period of retrospective analysis. The login screen identifies as MyChart for University of Iowa Health Care and allows for login using username and password, a link to sign up for access, or option for username or password recovery. Once logged in, the user has menu options for *Health* (eg, test results, health summary, current health issues, medications, allergies, and immunizations), *Visits* (viewing past, current, or future appointments and scheduling or canceling appointments), *Messaging* (inbox, ability to send and receive messages from health care team, and request for prescription refills), *Billing*, *Resources* (eg, medical library), and [User] *Profile*. Parents or legal guardians will see an option for accessing MyChart for their dependents. The Profile menu allows for various options for receiving notifications when new information (eg, diagnostic test results, medical documents, prescriptions, and bills) is available. Diagnostic test results can be accessed directly off the *Health* menu. For any given laboratory test, there is the ability to view past results (grouped into columns by date) and also graph/trend past results.

The institutional patient portal allows patient access to laboratory test results and radiologic imaging reports [41,42]. Authorized health care providers can manually *release* diagnostic test results to the patient portal once the results are finalized in the EHR; the results then immediately appear in the patient portal. Results that do not get manually released will *autorelease* based on a defined schedule. As mentioned above, patients can set up an email or other method for notifications when diagnostic test results become available in MyChart.

Common laboratory tests (eg, basic electrolyte and renal/kidney tests, cardiac troponin, complete blood count, coagulation tests, and urinalysis) autorelease following a 1–business day delay [41]. Radiologic imaging results and more sensitive laboratory tests (eg, sexually transmitted disease testing and genetic studies) autorelease with a 4–business day delay. A small number of other tests (eg, HIV testing and Huntington disease genetic studies) never release to the patient portal (*do not release* status). Other than HIV testing, diagnostic tests that do not release to the patient portal are very rarely ordered in the ED (the no-release status for HIV results in the patient portal allows for compliance with the State of Iowa law that requires patient counseling for reporting of positive HIV results) [41].

Data Retrieval and Analysis

There was a years' lag time between the retrospective period of diagnostic test ordering and our subsequent data analysis, allowing for longer-term assessment of patient portal access that may have occurred months after the ED patient encounter. This study was approved by the local institutional review board with a waiver for informed consent and was also part of a broader quality improvement initiative to improve patient use of the UIHC patient portal.

The study involved 2 primary measures. The first primary outcome involved analyzing patient portal activation rates for patients seen at the medical center ED who had one or more diagnostic tests performed. *Active* portal status was defined as those patients (or proxy) that registered their EHR portal account online with an activation code provided by the medical center. *Inactive* portal status was defined as those who had not used an activation or who did not elect to receive one. The second measure involved analysis of viewing patterns of diagnostic test results within the patient portal. Laboratory tests were divided into the categories of chemistry (including blood gas analysis), hematology (including blood count and coagulation testing), and microbiology (including blood and urine cultures along with more targeted infectious disease testing). Radiologic imaging was divided into the broader categories of computed tomography (CT), magnetic resonance imaging (MRI), and x-rays. Patient portal active status and viewing of diagnostic test results did not distinguish between patient or proxy access to the portal.

Epic Reporting Workbench (RWB) is a reporting tool within the EHR that can retrieve data based on specified query parameters [43]. The analysis for this study used the *Search Orders* template within RWB to look for laboratory test and radiologic imaging orders covering dates from October 1, 2016, to October 1, 2017, that were ordered during ED encounters and then completed (ie, excluded orders cancelled before being completed). The base functionality of the RWB search retrieves each specific order, patient identifier, and date/time of order. The search was enhanced to capture additional data fields associated with each order: patient gender, age, self-declared patient race, patient portal activity status (inactive or activated), when the diagnostic test result released to MyChart, whether the specific diagnostic test results were viewed in MyChart, and, if viewed, the date/time of MyChart viewing. Owing to the sheer number of diagnostic tests, the orderable menus for laboratory tests and radiologic imaging available during the retrospective time period were broken up into 20 separate search templates that were searched 2 weeks at a time to cover the year

XSL•FO
RenderX

of retrospective analysis. The outputs of the searches were downloaded to a spreadsheet format, with each row containing a unique test order and the additional data fields described above. The laboratory test and radiologic imaging data were separately combined, with subcategories assigned to specific diagnostic tests (eg, x-ray or magnetic resonance imaging). Spreadsheet analysis included pivot table analysis and was able to determine endpoints such as rates of portal activation and viewing rates of specific diagnostic tests associated with patient demographic features. The date/time stamps associated with the result release from the EHR to MyChart and patient access of results allowed for calculation of lag time between result release and viewing.

The categories listed for self-declared race were based on their primary response documented in the EHR. The groups included African American/black, Asian, Hispanic/Latino, other, and white. The category listed as *other* was for all patients that did not fall into any of the others. This included patients that indicated Native American, Alaskan, declined to answer, multiracial, Pacific Islander, Hawaiian, or unknown.

Statistical Analysis

The dataset for laboratory tests included 208,635 tests performed in the ED on 25,361 unique ED patients, of which 9482 (37.39%) had an active portal account. Note that the laboratory test dataset is limited only to those patients who had one or more laboratory tests performed in the ED. Analysis was performed separately for the total cohort (includes those who have not activated their patient portal account and thus could not view results) and for the subset limited only to those who had an active patient portal account. Analysis of the 2 populations helps address 2 variables that influence viewing of laboratory tests: (1) activation of the portal and (2) use of the portal once an account is activated. We conducted chi-square tests to compare view rates (viewed or not viewed) of laboratory tests by age category, gender, and self-declared race. As a total of 42 comparisons were made, the Bonferroni correction was used, and the significance level was set at .0011.

A similar approach was used for the radiologic imaging data. The cohort that had radiologic imaging studies performed in the ED included 23,504 studies on 14,455 unique patients, with 37.61% (5436) of these patients having an active patient portal. This imaging data only included patients who had one or more imaging studies ordered while in the ED. We conducted chi-square tests to compare view rates (viewed or not viewed) of imaging reports by age category, gender, and self-declared race. The Bonferroni correction was used, and the significance level was set at .0011.

Results

Rates of Patient Portal Activation and Viewing of Laboratory Tests by Patients Seen in the Emergency Department

During the retrospective analysis period, 208,635 tests on 25,361 unique patients were performed in the ED, of which 37.39% (n=9482) of these patients had an active patient portal account. In terms of the number of ED visits, 80.56% (n=20,430) unique patients had a single ED visit with laboratory testing, 16.04% (n=4069) had 2 or 3 ED visits, 3.16% (n=802) had 4 to 10 ED visits, and only 0.24% (n=60) had more than 10 ED visits. Activation rates were lower for those with only a single ED visit (7312/20,430, 35.79%) compared with either those with 2 to 3 ED visits (1770/4069, 43.50%; $P<.001$) or 4 or more ED visits (368/862, 42.7%; $P<.001$). Females (5546/13,149, 42.18%) were significantly more likely to activate their portal than males (3897/12,212, 31.91%; $P<.001$).

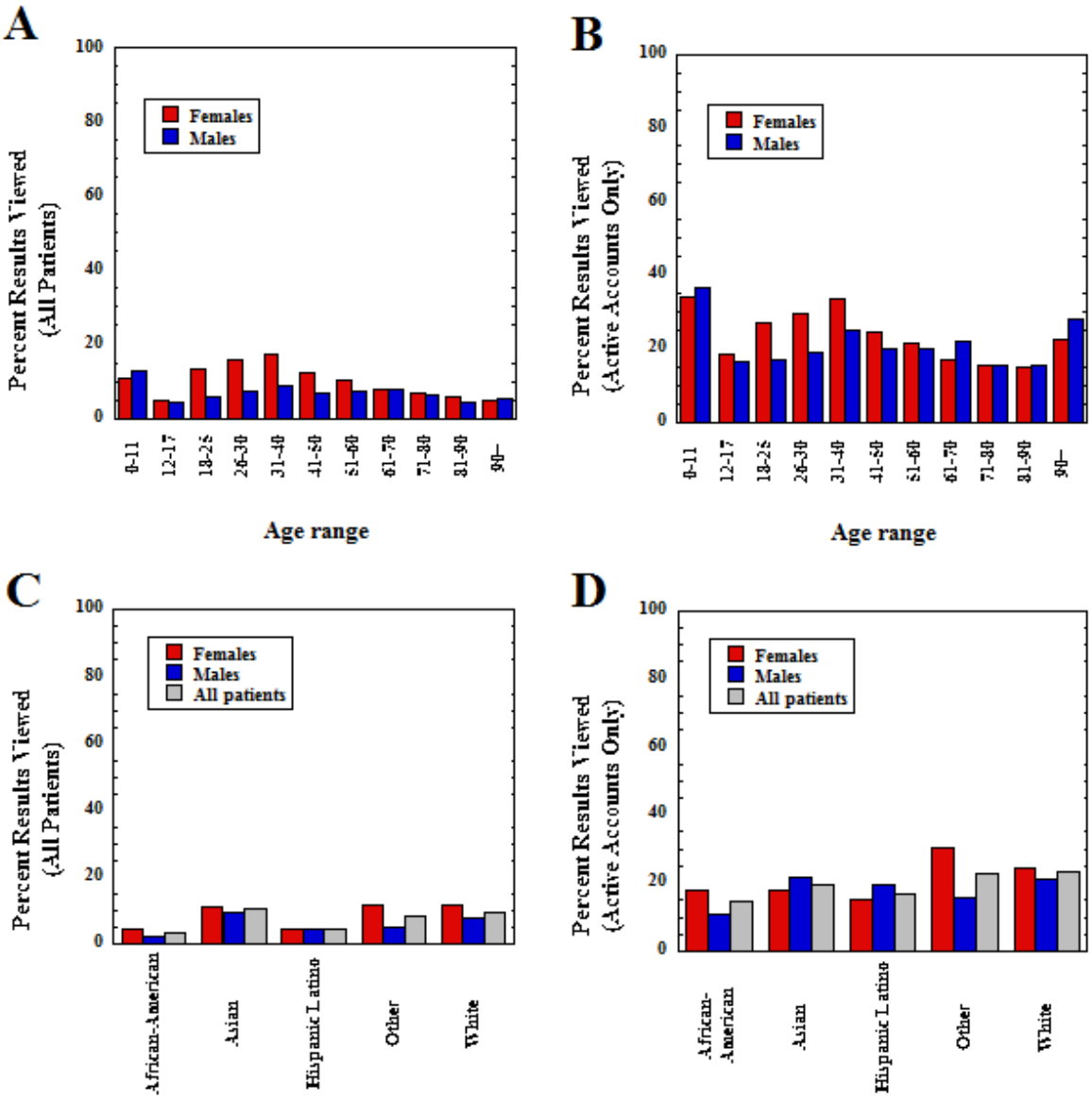
Activation rates were highest for those with self-declared race as Asian (262/451, 58.1%) and white (8155/20,637, 39.52%) and lower for African American/black (491/2254, 21.78%; $P<.001$ compared with white), Hispanic/Latino (333/1257, 26.49%; $P<.001$ compared with white), and other (241/762, 31.6%; $P<.001$ compared with white). Activation rates for patients aged 18 to 70 years were 41.61% (7593/18,246); all subgroups within that broad age range (18-25 years, 26-30 years, 31-40 years, 41-50 years, 51-60 years, and 61-70 years) were at least 39.1%. Activation rates were lower for those aged 0 to 11 years (437/1626, 26.88%), 12 to 17 years (215/1129, 19.04%), 71 to 80 years (785/2447, 32.08%), 81 to 90 years (385/1568, 24.55%), and older than 91 years (67/345, 19.4%; $P<.001$ for each subgroup compared with 18-70 years).

Figure 1 shows the rate of laboratory test viewing in the patient portal, broken down by age and gender, for all patients and limited to only those with an activated patient portal account. Overall, 8.91% of laboratory tests (18,573/208,635) ordered in the ED were viewed in the patient portal. Several broad trends were evident. In general, females viewed laboratory test results at higher rates (10.73%, n=11,049) than males (7.20%, n=7524; $P<.001$), with viewing rates by females approximately double or more than that of males in the 18 to 50 years age range, the years of peak viewing. Females viewed laboratory results at rates higher than males for every age subgroup between 18 and 90 years with the exception of 71 to 80 years ($P<.001$ for each comparison; $P=.10$ for 71 to 80 years). Viewing rates were similar between males and females in other age ranges and lowest overall in the 12 to 17 years category. For the patients that had active portal accounts, females viewed at rates significantly greater than males overall ($P<.001$) and for all age categories between 18 and 50 years ($P<.001$ for each subgroup in this age range).

Figure 1 also breaks down the laboratory viewing data by gender and self-reported race, separating into all patients and limited to those with an active portal account. One notable trend is that Hispanic/Latino and African American/black groups had the lowest viewing rates compared with the other race categories, which is especially evident in the overall patient graph ($P<.001$ for each, compared with white). Both white and Asian viewing rates approximate the overall rates for the ED population, reflecting that these 2 categories constitute a majority of the population in the ED. The differences between the race categories are less striking when restricted to not only those with active portal accounts but also those with significantly lower viewing rates for Hispanic/Latino and African American/black groups ($P<.001$ for each, compared with white). African American/black males had the lowest viewing rates of any subgroup.

XSL•FO
RenderX

Figure 1. View rates of emergency department laboratory tests and variation by age, gender (a and b), and self-declared race (c and d). Population includes all patients who had at least one laboratory test performed during period of retrospective analysis. Data based on 25,361 unique patients (2254 African American/black, 451 Asian, 1257 Hispanic/Latino, 762 other, and 20,637 white; 13,157 female and 12,204 male) and 208,635 tests (103,548 for female patients and 105,087 for male patients). Bars show the percentage of results viewed within each subcategory.



Multimedia Appendix 1 shows a breakdown of data by self-reported race and age. For the total population with laboratory testing ordered in the ED (25,361 unique patients), African American/black and Hispanic/Latino groups had lower laboratory test viewing rates in every age category except 26 to 30 years for Hispanic/Latino (P<.001 for each subgroup compared with a comparable white subgroup). Differences were less pronounced when examining only the subset with activated portal accounts (9482 unique patients). Nevertheless, African African/black individuals had lower view rates for all categories compared with white (P<.001). Hispanic/Latino individuals were lower than white in all age categories (P<.001) except 0 to 11 years and 18 to 25 years.

Table 1 shows the laboratory tests with the highest frequency of viewing that were ordered in the ED. Infectious disease testing [*Chlamydia trachomatis* polymerase chain reaction (PCR), *Neisseria gonorrhoeae* PCR, blood culture, and urine culture] and serum human chorionic gonadotropin (generally ordered for pregnancy testing) accounted for 5 of 10 most highly viewed tests. In contrast, examples of highly ordered tests with very low view rates (<5.0% overall) include arterial blood gas, chloride, and plasma ethanol. The broad categories of chemistry, hematology, and microbiology tests had similar view rates (Table 1).

XSL•FO
RenderX

Table 1. Viewing rates of laboratory tests by emergency department patients.

Laboratory category or test	Laboratory tests viewed	
	N	n (%)
Broad category		
Chemistry	120,047	10,499 (8.75)
Hematology	64,244	5804 (9.03)
Microbiology	24,330	2270 (9.33)
Specific laboratory test (at least 100 results for active portal users)		
Urine culture	729	106 (14.5)
D-dimer	1273	178 (13.98)
Chlamydia trachomatis polymerase reaction (PCR)	940	127 (13.5)
Blood culture	4250	570 (13.41)
Human chorionic gonadotropin, serum	665	88 (13.2)
Neisseria gonorrhoeae PCR	943	121 (12.8)
C-reactive protein	4679	542 (11.61)
Total protein	348	39 (11.2)
Complete blood count with differential	26,493	2913 (11.00)
Gamma-glutamyl transferase	432	47 (10.9)
Urinalysis	2976	311 (10.45)
Basic metabolic panel (“Chem 8”)	23,536	2414 (10.26)
Amylase	6052	614 (10.14)
Thyroid-stimulating hormone with reflex	4474	422 (9.43)
Bilirubin, total	884	78 (8.8)
Alkaline phosphatase	866	74 (8.5)
Hemoglobin, glycosylated (A_{1c})	659	56 (8.5)
Creatinine, plasma	567	47 (8.3)
Aspartate aminotransferase	1350	110 (8.14)
Drug of abuse screen—urine panel	4343	345 (7.94)

Rates of Patient Activation and Viewing of Radiologic Imaging Reports by Patients Who Had Imaging Performed While in the Emergency Department

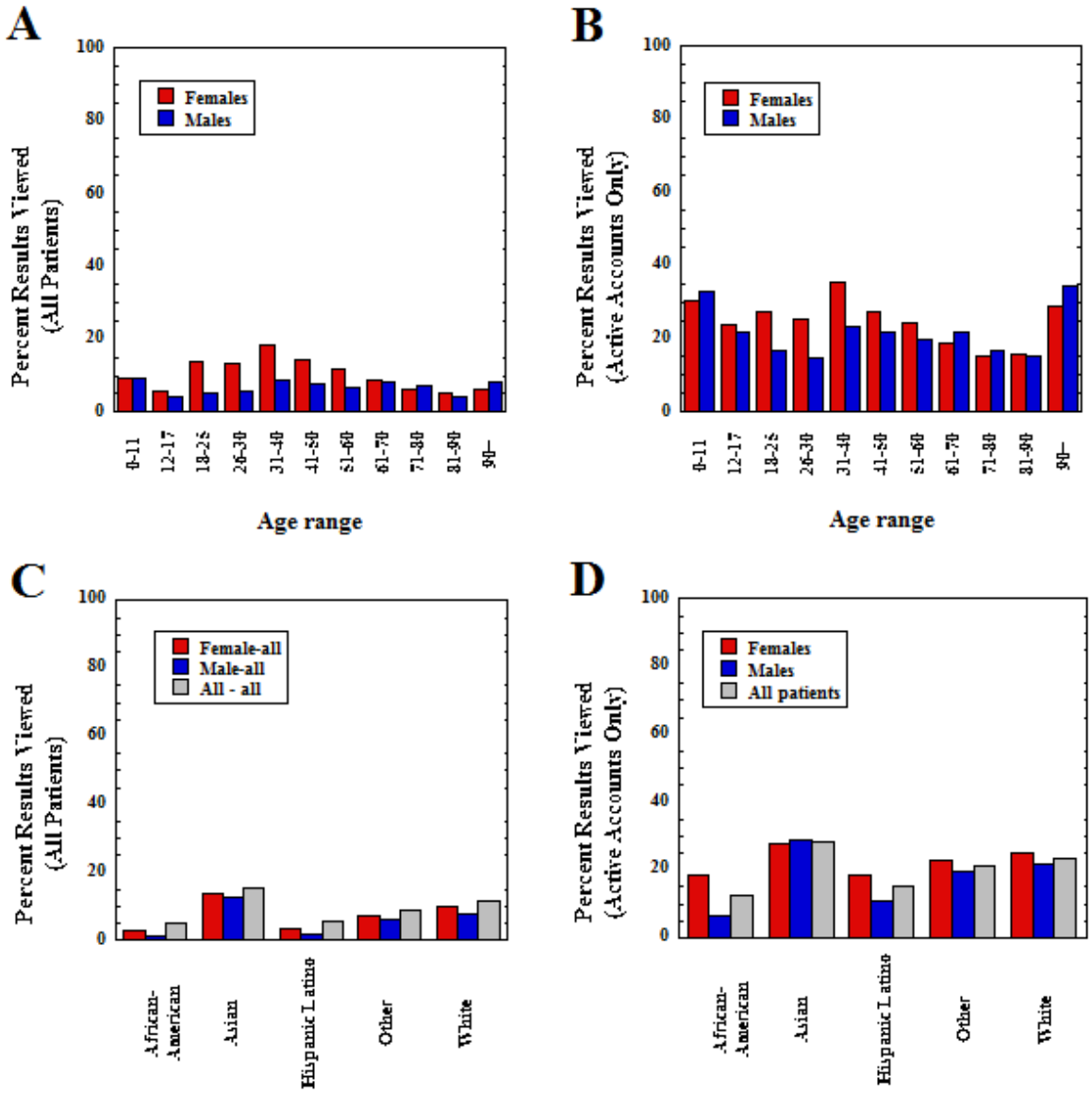
During the retrospective analysis period, the cohort that had radiologic imaging studies performed in the ED included 23,504 studies on 14,455 unique patients, with 37.61% (5436/14,455) of these patients having an active patient portal account. Figure 2 shows the rate of radiologic imaging report viewing in the patient portal, broken down by age and gender for all patients and limited to only those with an activated patient portal account. Overall, 8.97% (2019/22,504) of the reports from radiologic imaging studies ordered in the ED were viewed in the patient portal.

The overall pattern of viewing with respect to age and gender showed similar trends to those described above for laboratory testing. Highest rates of viewing were seen in the age ranges of 0 to 11 years, 18 to 50 years, and older than 90 years. Females viewed imaging reports at rates higher than males both overall (P<.001) and for every age subgroup (overall cohort) between 18 and 60 years (P<.01 for each comparison). For those with active portal accounts only, females viewed imaging reports at higher rates than males in all age categories between 18 and 60 years (P<.001 for all groups except P=.004 for 26-30 years, P=.03 for 41-50 years, and P=.03 for 51-60 years).

Figure 2 also breaks down the radiology report viewing data by gender and race, separating into all patients and limited to those with an active portal account. Trends were very similar to the laboratory result viewing data (Figure 1), with the lowest viewing rates in the Hispanic/Latino and African American/black groups, especially in the overall patient graph (P<.001 for each compared with white). Similar to laboratory results, the differences between the race categories are less striking when restricted to not only those with active portal accounts but also with significantly lower viewing rates for Hispanic/Latino and African American/black groups (P<.001 for African American/black compared with white and P=.003 for Hispanic/Latino compared with white). Also similar to laboratory results, African American/black males had the lowest viewing rates of any subgroup.

XSL•FO
RenderX

Figure 2. View rates of emergency department radiology tests and variation by age, gender (a and b), and self-declared race (c and d). Population includes all patients who had at least one radiologic imaging study performed during period of retrospective analysis. Data based on 14,455 unique patients (1274 African American/black, 220 Asian, 666 Hispanic/Latino, 407 other, and 11,888 white; 7079 female and 7376 male) and 23,504 radiology studies (11,202 for female patients and 12,302 for male patients). Bars show percent of results viewed within each subcategory.



Overall, MRI studies were reviewed at higher rates (329/2796, 11.77%) compared with CT scans (859/8957, 9.59%) and x-rays (911/11,750, 7.75%; $P<.001$ for MRI compared with either CT scan or x-ray for all ED patients). Table 2 shows the most viewed radiology studies with highest rates of viewing that were ordered in the ED, limited to studies with at least 50 orders in the retrospective analysis period. MRI studies (brain and spine imaging) and CT scans (abdomen/pelvis, chest, and head/neck) constituted 8 of the 10 tests in the patient portal with the highest frequency of viewing, with some studies viewed by over 30% (eg, MRI cervical spine with contrast: 25/73, 34%) of those with active portal accounts.

XSL•FO
RenderX

Table 2. Viewing rates of radiologic imaging reports by emergency department patients.

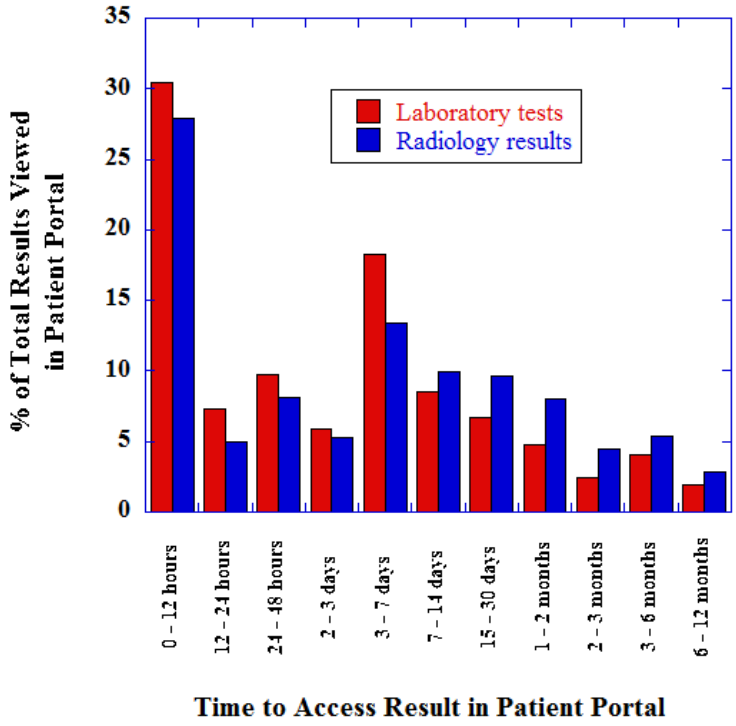
Radiologic imaging category or test	Imaging reports viewed	
	N	n (%)
Broad category		
Magnetic resonance imaging (MRI)	2796	329 (11.77)
Computed tomography (CT) scan	8957	859 (9.59)
X-ray study	11,750	911 (7.75)
Specific radiologic imaging study (at least 50 studies for all patients)		
MRI cervical spine with contrast	136	27 (19.9)
CT abdomen and pelvis without contrast	837	131 (15.7)
MRI brain—axial T2	188	27 (14.4)
CT abdomen and pelvis with contrast	2106	299 (14.19)
CT chest with contrast	175	22 (12.6)
MRI brain with contrast	1774	215 (12.12)
X-ray chest—anteroposterior and lateral	6418	644 (10.03)
MRI cervical spine without contrast	240	21 (8.8)
X-ray C-spine—anteroposterior and lateral	186	15 (8.1)
MRI brain without contrast	274	21 (7.7)
X-Ray abdomen— anteroposterior supine	468	35 (7.5)
CT brain without contrast	5137	360 (7.01)
X-ray chest anteroposterior/posteroanterior	3441	164 (4.77)
X-ray pelvis—anteroposterior	1109	46 (4.14)

Time to Access Results in Patient Portal

Figure 3 shows how quickly patients accessed laboratory results or radiology reports after release into the patient portal. Only data from results that were accessed in the patient portal are included in Figure 3. There was a wide variability in how fast the results were accessed: laboratory results (mean 17.1, SD 42.7 days; median 2.3 days; minimum 0.9 min, maximum 359 days) and radiology reports (mean 24.0, SD 48.3 days; median 3.7 days; minimum 0.9 min, maximum 336 days). Compared with radiology results, laboratory tests tended to be more frequently viewed in time categories less than 1 week (P<.001). Nearly half of all results accessed in the patient portal are reviewed within 72 hours of release to the patient portal (laboratory results: 9904/18,573, 53.32%; radiology reports: 971/2019, 48.1%). On the other extreme, laboratory results and radiology reports viewed after 2 weeks constituted 19.93% (3701/18,573) and 971/2019, 48.1%, respectively, of the total views.

XSL•FO
RenderX

Figure 3. Time in which patient (or proxy) accesses laboratory test or radiologic imaging results after release to the patient portal. The data are divided into various time categories, summarizing the percentage of total results viewed in those categories.



Discussion

Principal Findings

In this study, we show that activation and utilization of an EHR portal for viewing of diagnostic test results by patients at an academic medical center ED was the highest for those who were female, who were aged 0 to 11 years and 18 to 60 years, and with self-reported race as Caucasian or Asian. The lowest rates were for those who were teenagers, aged older than 81 years, African American/black, and Hispanic/Latino.

Patient portals are an increasingly popular feature for institutions using EHRs, offering a potentially more convenient way for patients to access their health information, coordinate appointments, and communicate with their health care team [5-11]. Previous studies have shown that patients who actively use portals are more likely to be female, be young to middle-aged adults, be English speaking, be Caucasian, have better health care literacy, and be in better health than nonusers [8,9,16,23,26-30,44]. Other studies have shown disparities in the use of EHR portals by underrepresented minorities [45-47]. There are complex individual factors that influence patient portal adoption [48]. The match between the design of the portal and patient expectations influences adoption and continued use [49], and some institutions have extensively incorporated stakeholder input into the portal design [50]. An example of incorporating user input to improve the adoption and use of MyChart is a project at the Johns Hopkins School of Medicine that trained volunteers to educate adolescent and young adult patients about the patient portal and to then facilitate enrollment and, for those interested, downloads of the MyChart mobile app [51].

In this study, we focused on the use of the patient portal for accessing diagnostic test results in an ED setting, comparing overall view rates (impacted by portal *nonusers* in addition to those with active accounts) with view rates among active portal users. The trends between laboratory results and radiology reports are very similar and will be discussed together. The highest use for both types of results occurred in the following groups: females; ages 0 to 11 years and 26 to 60 years; and white, Asian, and other self-declared race categories. Disparities between African American/black and Hispanic/Latino groups were more pronounced when viewing the overall patient population compared with just the subgroup that has an active portal account, demonstrating that low activation rates are a significant factor driving overall low viewing rates for African American/black and Hispanic/Latino patients in our study.

There are multiple approaches to reducing disparities in patient portal adoption. One positive step has been the development of multilingual patient portals, as described for a California health system serving a predominantly Spanish-speaking population [52]. Other factors that may influence portal usage include lack of broadband internet access, ownership of smartphones, and presence of family members or friends to assist with portal usage [33,45-47]. Initiatives to promote portal activation and usage should take into account these potential barriers. Future studies can also focus on the association between particular diseases/diagnoses and the use of the patient portal in the ED and other settings.

The rates of patient portal viewing of diagnostic tests in the ED in this study are considerably lower than in our previous study, which focused on outpatient clinics [42]. For outpatients, portal account activation rates at UIHC were 39.9% for females and

XSL•FO
RenderX

31.9% for males; the overall viewing of outpatient diagnostic tests results was nearly 40%, with some subgroups exceeding 55%. Very little has been published about how ED patients use portals, and our institution is probably similar to many others in having the broadest functionality of the patient portal for the ambulatory care setting. Efforts to promote patient portals in the ED can focus on the benefits of accessing information such as diagnostic testing and discharge summaries. This may help patients better understand their conditions as well as bring information and questions to future health care encounters. A visit in the ED may also be someone's first exposure to a patient portal that can be used in other health care settings.

There has been increasing research related to patient portal usage for the pediatric population [53,54]. In this study, viewing of diagnostic test results in the 0 to 11 years patient age range (when parents/guardians have unrestricted proxy access) was similar to that in the 18 to 60 years range, with the exception that the access rates did not vary based on the gender of the child. The dataset did not capture which parent or guardian accessed the child's results, so it is unknown if there are gender differences in proxy access or how often 2 parents/guardians view results together. Age ranges with low overall access of diagnostic test results include 12 to 17 years and older than 70 years. The institutional policy at UIHC limits proxy access to the patient portal for children aged 12 to 17 years; thus, efforts to increase portal usage in this age range have to be age appropriate. Another study has shown that older adults may be unaware of what medical information they can access via a patient portal, potentially limiting their enthusiasm for the use of the patient portal [55]. They may also be unaware of the different features available and the variety of ways that they can access information via patient portals. Training and assistance from family and friends may be helpful in overcoming resistance and lack of experience or comfort with the use of new technology such as patient portals [16,29,35,55].

Finally, this study shows that ED patient access of the diagnostic test results spans a wide time range, with some patients accessing results within minutes of availability in the patient portal (likely because of quick access to the patient portal following email or other notification that a diagnostic test result was available) and others only months after the ED encounter. Median view rates were between 2 and 4 days for laboratory and radiology results. There was a large spike that occurred in the 3 to 7 days range. This pattern is different from what we observed with viewing of results by outpatients at UIHC, where the most prominent spike in viewing is within 24 hours of availability in the patient portal, with a steady tail in viewing patterns after 24 hours [41,42]. For ED patients, there are a variety of factors that could influence viewing patterns. There could be delays from patients recovering from the cause of the visit before going and viewing results. If the patient is admitted to an inpatient unit from the ED, access to results would be influenced by factors such as internet access within the hospital and assistance from family and friends. Results from the testing performed in the ED could also become important later as patients have follow-up visits with outpatient providers. Finally, the nature of the patient's illness likely has an impact on the use of a patient portal.

Limitations

The limitations of this study include that data were obtained only from a single academic medical center ED with has a predominantly white patient population. Activation and access data for the patient portal also cannot capture the impact these data have on patient experience, especially as many patients seen at the institutional ED (which serves as the highest-complexity, level 1 trauma center in a wide geographic region) receive primary care at other health care institutions not affiliated with UIHC. In addition, as mentioned above, the data did not distinguish between patient or proxy access or the identity (eg, gender and age) of the person doing proxy access.

Conclusions

We found that patient portal access of ED diagnostic test results was highest in the following populations: females, those with self-reported race as Caucasian or Asian, and those aged 0 to 11 years and 18 to 60 years. Groups with the lowest access rates included African American males, Hispanic/Latino individuals, teenagers, and those aged 81 years or older. Differences between groups in terms of diagnostic test access were less striking when looking at only those with activated patient portal accounts, suggesting that getting patients to activate a portal account is an important barrier. It is also important to note that even groups with the highest access rate viewed diagnostic test results at less than 20%, leaving much opportunity for improvement. More research should be done to ascertain the strategies for increasing patient portal usage in the ED. Data on patient portal usage can guide education effects to enhance patient engagement and minimize disparities between patient groups.

Authors' Contributions

MDK had full access to all the data in the study, takes responsibility for the integrity of the data, and is accountable for the accuracy of the analysis. BF wrote the first draft of the manuscript and performed the majority of data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Breakdown of data by self-reported race and age.

[PDF File (Adobe PDF File), 357KB - jmir_v21i6e13791_app1.pdf]

XSL•FO
RenderX

References

1. Adler-Milstein J, DesRoches CM, Kralovec P, Foster G, Worzala C, Charles D, et al. Electronic health record adoption in US hospitals: progress continues, but challenges persist. Health Aff (Millwood) 2015 Dec;34(12):2174-2180. [doi: 10.1377/hlthaff.2015.0992] [Medline: 26561387]
2. DesRoches CM, Charles D, Furukawa MF, Joshi MS, Kralovec P, Mostashari F, et al. Adoption of electronic health records grows rapidly, but fewer than half of US hospitals had at least a basic system in 2012. Health Aff (Millwood) 2013 Aug;32(8):1478-1485. [doi: 10.1377/hlthaff.2013.0308] [Medline: 23840052]
3. Heisey-Grove D, Danehy L, Consolazio M, Lynch K, Mostashari F. A national study of challenges to electronic health record adoption and meaningful use. Med Care 2014 Feb;52(2):144-148. [doi: 10.1097/MLR.0000000000000038] [Medline: 24309669]
4. Kern LM, Edwards A, Kaushal R, HITEC Investigators. The meaningful use of electronic health records and health care utilization. Am J Med Qual 2016 Dec;31(4):301-307. [doi: 10.1177/1062860615572439] [Medline: 25712134]
5. Eschler J, Meas P, Lozano P, McClure J, Ralston J, Pratt W. Integrating the patient portal into the health management work ecosystem: user acceptance of a novel prototype. AMIA Annu Symp Proc 2016;2016:541-550 [FREE Full text] [Medline: 28269850]
6. Fix GM, Hogan TP, Amante DJ, McInnes DK, Nazi KM, Simon SR. Encouraging patient portal use in the patient-centered medical home: three stakeholder perspectives. J Med Internet Res 2016 Nov 22;18(11):e308 [FREE Full text] [doi: 10.2196/jmir.6488] [Medline: 27876686]
7. Gheorghiu B, Hagens S. Use and maturity of electronic patient portals. Stud Health Technol Inform 2017;234:136-141. [Medline: 28186030]
8. Griffin A, Skinner A, Thornhill J, Weinberger M. Patient portals: who uses them? What features do they use? And do they reduce hospital readmissions? Appl Clin Inform 2016;7(2):489-501. [doi: 10.4338/ACI-2016-01-RA-0003] [Medline: 27437056]
9. Nazi K, Woods S, Woods SS. MyHealtheVet PHR: a description of users and patient portal use. AMIA Annu Symp Proc 2008 Nov 06;6:1182. [Medline: 18999142]
10. Pensold P, Schneider G, Aguduri L, Brandner A, Bronsch T, Schreiweis B, et al. Implementing the patient portal of the PEHR in the metropolitan region Rhine-Neckar-experiences with Liferay. Stud Health Technol Inform 2016;228:105-109. [Medline: 27577351]
11. Wolff JL, Berger A, Clarke D, Green JA, Stametz R, Yule C, et al. Patients, care partners, and shared access to the patient portal: online practices at an integrated health system. J Am Med Inform Assoc 2016 Nov;23(6):1150-1158. [doi: 10.1093/jamia/ocw025] [Medline: 27026614]
12. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
13. Lyles CR, Fruchterman J, Youdelman M, Schillinger D. Legal, practical, and ethical considerations for making online patient portals accessible for all. Am J Public Health 2017 Dec;107(10):1608-1611. [doi: 10.2105/AJPH.2017.303933] [Medline: 28817324]
14. Mennemeyer ST, Menachemi N, Rahurkar S, Ford EW. Impact of the HITECH Act on physicians' adoption of electronic health records. J Am Med Inform Assoc 2016 Mar;23(2):375-379. [doi: 10.1093/jamia/ocv103] [Medline: 26228764]
15. Brown SM, Bell SK, Roche SD, Dente E, Mueller A, Kim T, et al. Preferences of current and potential patients and family members regarding implementation of electronic communication portals in intensive care units. Ann Am Thorac Soc 2016 Mar;13(3):391-400. [doi: 10.1513/AnnalsATS.201509-638OC] [Medline: 26700656]
16. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito DA. Patient portals as a tool for health care engagement: a mixed-method study of older adults with varying levels of health literacy and prior patient portal use. J Med Internet Res 2017 Mar 30;19(3):e99 [FREE Full text] [doi: 10.2196/jmir.7099] [Medline: 28360022]
17. Neuner J, Fedders M, Caravella M, Bradford L, Schapira M. Meaningful use and the patient portal: patient enrollment, use, and satisfaction with patient portals at a later-adopting center. Am J Med Qual 2015;30(2):105-113. [doi: 10.1177/1062860614523488] [Medline: 24563085]
18. Turvey C, Klein D, Fix G, Hogan TP, Woods S, Simon SR, et al. Blue Button use by patients to access and share health record information using the Department of Veterans Affairs' online patient portal. J Am Med Inform Assoc 2014;21(4):657-663. [doi: 10.1136/amiajnl-2014-002723] [Medline: 24740865]
19. Ammenwerth E, Schnell-Inderst P, Hoerbst A. The impact of electronic patient portals on patient care: a systematic review of controlled trials. J Med Internet Res 2012;14(6):e162 [FREE Full text] [doi: 10.2196/jmir.2238] [Medline: 23183044]
20. Reicher JJ, Reicher MA. Implementation of Certified EHR, Patient Portal, and "Direct" Messaging Technology in a Radiology Environment Enhances Communication of Radiology Results to Both Referring Physicians and Patients. J Digit Imaging 2016 Dec;29(3):337-340 [FREE Full text] [doi: 10.1007/s10278-015-9845-x] [Medline: 26588906]
21. Sorondo B, Allen A, Fathima S, Bayleran J, Sabbagh I. Patient portal as a tool for enhancing patient experience and improving quality of care in primary care practices. EGEMS (Wash DC) 2016;4(1):1262 [FREE Full text] [doi: 10.13063/2327-9214.1262] [Medline: 28203611]

XSL•FO
RenderX

22. Moll J, Rexhepi H, Cajander Å, Grünloh C, Huvila I, Hägglund M, et al. Patients' experiences of accessing their electronic health records: national patient survey in Sweden. J Med Internet Res 2018 Dec 01;20(11):e278 [FREE Full text] [doi: 10.2196/jmir.9492] [Medline: 30389647]
23. Curtis J, Cheng S, Rose K, Tsai O. Promoting adoption, usability, and research for personal health records in Canada: the MyChart experience. Healthc Manage Forum 2011;24(3):149-154. [doi: 10.1016/j.hcmf.2011.07.004] [Medline: 22165576]
24. Nazi KM, Hogan TP, McInnes DK, Woods SS, Graham G. Evaluating patient access to electronic health records: results from a survey of veterans. Med Care 2013 Mar;51(3 Suppl 1):S52-S56. [doi: 10.1097/MLR.0b013e31827808db] [Medline: 23407012]
25. Redelmeier DA, Kraus NC. Patterns in patient access and utilization of online medical records: analysis of MyChart. J Med Internet Res 2018 Feb 06;20(2):e43 [FREE Full text] [doi: 10.2196/jmir.8372] [Medline: 29410386]
26. Gerber DE, Laccetti AL, Chen B, Yan J, Cai J, Gates S, et al. Predictors and intensity of online access to electronic medical records among patients with cancer. J Oncol Pract 2014 Sep;10(5):e307-e312 [FREE Full text] [doi: 10.1200/JOP.2013.001347] [Medline: 25006222]
27. McInnes DK, Shimada SL, Midboe AM, Nazi KM, Zhao S, Wu J, et al. Patient use of electronic prescription refill and secure messaging and its association with undetectable HIV viral load: a retrospective cohort study. J Med Internet Res 2017 Feb 15;19(2):e34 [FREE Full text] [doi: 10.2196/jmir.6932] [Medline: 28202428]
28. Miller H, Vandenbosch B, Ivanov D, Black P. Determinants of personal health record use: a large population study at Cleveland Clinic. J Healthc Inf Manag 2007;21(3):44-48. [Medline: 19195293]
29. Nahm E, Sagherian K, Zhu S. Use of patient portals in older adults: a comparison of three samples. Stud Health Technol Inform 2016;225:354-358. [Medline: 27332221]
30. Wallace LS, Angier H, Huguet N, Gaudino JA, Krist A, Dearing M, et al. Patterns of electronic portal use among vulnerable patients in a nationwide practice-based research network: from the OCHIN Practice-based Research Network (PBRN). J Am Board Fam Med 2016 Oct;29(5):592-603 [FREE Full text] [doi: 10.3122/jabfm.2016.05.160046] [Medline: 27613792]
31. Jhamb M, Cavanaugh KL, Bian A, Chen G, Ikizler TA, Unruh ML, et al. Disparities in electronic health record patient portal use in nephrology clinics. Clin J Am Soc Nephrol 2015 Nov 06;10(11):2013-2022 [FREE Full text] [doi: 10.2215/CJN.01640215] [Medline: 26493242]
32. Nazi KM. Veterans' voices: use of the American Customer Satisfaction Index (ACSI) Survey to identify My HealtheVet personal health record users' characteristics, needs, and preferences. J Am Med Inform Assoc 2010;17(2):203-211 [FREE Full text] [doi: 10.1136/jamia.2009.000240] [Medline: 20190065]
33. Trubitt M, Alozie O, Shokar G, Flores S, Shokar NK. Patterns and correlates of internet use, cell phone use, and attitudes toward patient portals among a predominantly Mexican-American clinic population. Telemed J E Health 2018 Feb 02;24(11):861-869. [doi: 10.1089/tmj.2017.0266] [Medline: 29393769]
34. Walker DM, Menser T, Yen P, McAlearney AS. Optimizing the user experience: identifying opportunities to improve use of an inpatient portal. Appl Clin Inform 2018 Dec;9(1):105-113. [doi: 10.1055/s-0037-1621732] [Medline: 29444536]
35. Wildenbos G, Peute L, Jaspers M. Facilitators and barriers of electronic health record patient portal adoption by older adults: a literature study. Stud Health Technol Inform 2017;235:308-312. [Medline: 28423804]
36. Huerta T, Fareed N, Hefner JL, Sieck CJ, Swoboda C, Taylor R, et al. Patient engagement as measured by inpatient portal use: methodology for log file analysis. J Med Internet Res 2019 Mar 25;21(3):e10957 [FREE Full text] [doi: 10.2196/10957] [Medline: 30907733]
37. Kelly MM, Dean SM, Carayon P, Wetterneck TB, Hoonakker PL. Healthcare team perceptions of a portal for parents of hospitalized children before and after implementation. Appl Clin Inform 2017 Dec 15;8(1):265-278 [FREE Full text] [doi: 10.4338/ACI-2016-11-RA-0194] [Medline: 28293685]
38. Masterson CR, Prey J, Ryan B, Alarcon I, Qian M, Bakken S, et al. Engaging hospitalized patients in clinical care: study protocol for a pragmatic randomized controlled trial. Contemp Clin Trials 2016 Jan 18;47:165-171. [doi: 10.1016/j.cct.2016.01.005] [Medline: 26795675]
39. Winstanley EL, Burtchin M, Zhang Y, Campbell P, Pahl J, Beck S, et al. Inpatient experiences with MyChart bedside. Telemed J E Health 2017 Dec;23(8):691-693 [FREE Full text] [doi: 10.1089/tmj.2016.0132] [Medline: 28328391]
40. Callen J, Giardina TD, Singh H, Li L, Paoloni R, Georgiou A, et al. Emergency physicians' views of direct notification of laboratory and radiology results to patients using the Internet: a multisite survey. J Med Internet Res 2015 Mar 04;17(3):e60 [FREE Full text] [doi: 10.2196/jmir.3721] [Medline: 25739322]
41. Krasowski MD, Grieme CV, Cassady B, Dreyer NR, Wanat KA, Hightower M, et al. Variation in results release and patient portal access to diagnostic test results at an academic medical center. J Pathol Inform 2017;8:45 [FREE Full text] [doi: 10.4103/jpi.jpi_53_17] [Medline: 29226008]
42. Oest SE, Hightower M, Krasowski MD. Activation and utilization of an electronic health record patient portal at an academic medical center-impact of patient demographics and geographic location. Acad Pathol 2018;5:2374289518797573 [FREE Full text] [doi: 10.1177/2374289518797573] [Medline: 30302394]
43. Krasowski MD, Chudzik D, Dolezal A, Steussy B, Gailey MP, Koch B, et al. Promoting improved utilization of laboratory testing through changes in an electronic medical record: experience at an academic medical center. BMC Med Inform Decis Mak 2015 Feb 22;15:11 [FREE Full text] [doi: 10.1186/s12911-015-0137-7] [Medline: 25880934]

XSL•FO
RenderX

44. Hoogenbosch B, Postma J, de Man-van Ginkel JM, Tiemessen NA, van Delden JJ, van Os-Medendorp H. Use and the users of a patient portal: cross-sectional study. J Med Internet Res 2018 Sep 17;20(9):e262 [FREE Full text] [doi: 10.2196/jmir.9418] [Medline: 30224334]
45. Ancker J, Hafeez B, Kaushal R. Socioeconomic disparities in adoption of personal health records over time. Am J Manag Care 2016 Dec;22(8):539-540 [FREE Full text] [Medline: 27541700]
46. Mook PJ, Trickey AW, Krakowski KE, Majors S, Theiss MA, Fant C, et al. Exploration of portal activation by patients in a healthcare system. Comput Inform Nurs 2018 Jan;36(1):18-26. [doi: 10.1097/CIN.0000000000000392] [Medline: 29049084]
47. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011 May 1;18(3):318-321 [FREE Full text] [doi: 10.1136/jamia.2010.006015] [Medline: 21262921]
48. Tavares J, Oliveira T. New integrated model approach to understand the factors that drive electronic health record portal adoption: cross-sectional national survey. J Med Internet Res 2018 Nov 19;20(11):e11032 [FREE Full text] [doi: 10.2196/11032] [Medline: 30455169]
49. Van den Bulck SA, Hermens R, Slegers K, Vandenberghe B, Goderis G, Vankrunkelsven P. Designing a patient portal for patient-centered care: cross-sectional survey. J Med Internet Res 2018 Oct 01;20(10):e269 [FREE Full text] [doi: 10.2196/jmir.9497] [Medline: 30287416]
50. Kildea J, Battista J, Cabral B, Hendren L, Herrera D, Hijal T, et al. Design and development of a person-centered patient portal using participatory stakeholder co-design. J Med Internet Res 2019 Feb 11;21(2):e11371 [FREE Full text] [doi: 10.2196/11371] [Medline: 30741643]
51. Ramsey A, Lanzo E, Huston-Paterson H, Tomaszewski K, Trent M. Increasing patient portal usage: preliminary outcomes from the MyChart genius project. J Adolesc Health 2018;62:29-35. [doi: 10.1016/j.jadohealth.2017.08.029] [Medline: 29169768]
52. Ochoa A, Kitayama K, Uijtdehaage S, Vermillion M, Eaton M, Carpio F, et al. Patient and provider perspectives on the potential value and use of a bilingual online patient portal in a Spanish-speaking safety-net population. J Am Med Inform Assoc 2017 Nov 01;24(6):1160-1164. [doi: 10.1093/jamia/ocx040] [Medline: 28460130]
53. Bush RA, Connelly CD, Fuller M, Pérez A. Implementation of the integrated electronic patient portal in the pediatric population: a systematic review. Telemed J E Health 2015 Aug 10;22(2):144-152 [FREE Full text] [doi: 10.1089/tmj.2015.0033] [Medline: 26258289]
54. Kidwell KM, Peugh J, Westcott E, Nwankwo C, Britto MT, Quinn CT, et al. Acceptability and feasibility of a disease-specific patient portal in adolescents with sickle cell disease. J Pediatr Hematol Oncol 2018 Oct 18 Epub ahead of print. [doi: 10.1097/MPH.0000000000001341] [Medline: 30339657]
55. Latulipe C, Quandt SA, Melius KA, Bertoni A, Miller DP, Smith D, et al. Insights into older adult patient concerns around the caregiver proxy portal use: qualitative interview study. J Med Internet Res 2018 Nov 02;20(11):e10524 [FREE Full text] [doi: 10.2196/10524] [Medline: 30389654]

Abbreviations

CT: computed tomography
ED: emergency department
EHR: electronic health record
HITECH: Health Information Technology for Economic and Clinical Health Act
MRI: magnetic resonance imaging
PCR: polymerase chain reaction
RWB: Reporting Workbench
UIHC: University of Iowa Hospitals and Clinics

Edited by G Eysenbach; submitted 22.02.19; peer-reviewed by S Connolly, N Fareed, K Colorafi, J Gee, R Sadeghi; comments to author 20.03.19; revised version received 14.05.19; accepted 04.06.19; published 28.06.19.

Please cite as:
Foster B, Krasowski MD
The Use of an Electronic Health Record Patient Portal to Access Diagnostic Test Results by Emergency Patients at an Academic Medical Center: Retrospective Study
J Med Internet Res 2019;21(6):e13791
URL: http://www.jmir.org/2019/6/e13791/
doi:10.2196/13791
PMID:31254335

XSL•FO
RenderX

©Brody Foster, Matthew David Krasowski. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development and Validation of a Personalized Social Media Platform–Based HIV Incidence Risk Assessment Tool for Men Who Have Sex With Men in China

Ke Yun[1,2,3,4*], PhD; Junjie Xu[1,2,3,4*], PhD; Sequoia Leuba[5], MPH, MD; Yunyu Zhu[6], BSc; Jing Zhang[1,2,3,4], PhD; Zhenxing Chu[1,2,3,4], MPH; Wenqing Geng[1,2,3,4], MPH; Yongjun Jiang[1,2,3,4], MPH; Hong Shang[1,2,3,4], PhD

[1]Key Laboratory of AIDS Immunology of National Health Commission of the People's Republic of China, Department of Laboratory Medicine, The First Affiliated Hospital of China Medical University, Shenyang, China

[2]Key Laboratory of AIDS Immunology of Liaoning Province, The First Affiliated Hospital of China Medical University, Shenyang, China

[3]Key Laboratory of AIDS Immunology, Chinese Academy of Medical Sciences, Shenyang, China

[4]Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Hangzhou, China

[5]Department of Epidemiology, University of North Carolina, Chapel Hill, American Samoa

[6]Hebei Yuanqiao Information Technology Co, Ltd, Shijiazhuang, China

[*]these authors contributed equally

Corresponding Author:
Hong Shang, PhD
Key Laboratory of AIDS Immunology of National Health Commission of the People's Republic of China
Department of Laboratory Medicine
The First Affiliated Hospital of China Medical University
No 155, Nanjing North Street, Heping District
Shenyang,
China
Phone: 86 8328 2634
Email: hongshang100@hotmail.com

Abstract

Background: Personalized risk assessments can help medical providers determine targeted populations for counseling and risk reduction interventions.

Objective: The objective of this study was to develop a social media platform–based HIV risk prediction tool for men who have sex with men (MSM) in China based on an independent MSM cohort to help medical providers determine target populations for counseling and risk reduction treatments.

Methods: A prospective cohort of MSM from Shenyang, China, followed from 2009 to 2016, was used to develop and validate the prediction model. The eligible MSM were randomly assigned to the training and validation dataset, and Cox proportional hazards regression modeling was conducted using predictors for HIV seroconversion selected by the training dataset. Discrimination and calibration were performed, and the related nomogram and social media platform–based HIV risk assessment tool were constructed.

Results: The characteristics of the sample between the training dataset and the validation dataset were similar. The risk prediction model identified the following predictors for HIV seroconversion: the main venue used to find male sexual partners, had condomless receptive or insertive anal intercourse, and used rush poppers. The model was well calibrated. The bootstrap C-index was 0.75 (95% CI 0.65-0.85) in the training dataset, and 0.60 (95% CI 0.45-0.74) in the validation dataset. The calibration plots showed good agreement between predicted risk and the actual proportion of no HIV infection in both the training and validation datasets. Nomogram and WeChat-based HIV incidence risk assessment tools for MSM were developed.

Conclusions: This social media platform–based HIV infection risk prediction tool can be distributed easily, improve awareness of personal HIV infection risk, and stratify the MSM population based on HIV risk, thus informing targeted interventions for MSM at greatest risk for HIV infection.

(J Med Internet Res 2019;21(6):e13475) doi:10.2196/13475

XSL•FO
RenderX

KEYWORDS

HIV; risk prediction; social media; men who have sex with men; China

Introduction

Background

A significant global challenge to HIV prevention and control is the HIV epidemic concentrated among men who have sex with men (MSM) [1]. In China, the annual reported incident HIV/AIDS cases among MSM is very high. In 2015, the HIV prevalence among MSM was 8%, and the HIV incidence among MSM was 5.61 cases per 100 person-years [2]. To end the global HIV epidemic by 2030, the Joint United Nations Programme on HIV and AIDS (UNAIDS) developed the 90-90-90 strategy, and the first "90" goal is to have 90% of HIV-infected people know their HIV serostatus by 2020 [3]. However, an estimated 50% of HIV-infected Chinese MSM do not know their serostatus [4], and only 60.5% of Chinese MSM have ever been tested for HIV in their lifetime [5]. This low HIV testing rate among Chinese MSM could be caused by an inaccurate evaluation of their personal HIV infection risk. A study examining perceived risk and HIV acquisition found that over 50% of participants diagnosed with HIV in the study thought their lifelong risk of HIV infection was very low or none [6].

Given that the risk of HIV infection among MSM varies, identifying MSM who are at higher risk of HIV seroconversion can improve targeted prevention interventions, which include HIV screening and preexposure prophylaxis (PrEP) [7]. In recent years, prospective cohort studies have examined HIV seroconversion predictors among MSM, including multiple sexual partners, condomless anal intercourse, sexually transmitted infections, and rush poppers usage [8]. Rush poppers are an inhalable mixture of various nitrites that are frequently used among the MSM community as they can relax smooth muscle, expand peripheral blood vessels, and relieve pain during anal sex, and thus increase sexual pleasure [8]. Insight into these predictors can lead to more accurate predictions of HIV seroconversion risk among MSM. Several risk evaluation models specific to MSM have been developed to quantify this risk including the Denver model [9], the University of North Carolina at Malawi Risk Screening Score model [10], and the San Diego Early Test (SDET) score model [11]. In addition, prediction models specific to Chinese MSM have been developed to estimate HIV infection risk [12,13]. However, all the previous models mentioned were developed using cross-sectional survey data, and thus could not predict longitudinal HIV seroconversion risk. The Menza score is a risk prediction model that could predict the risk of longitudinal HIV seroconversion because it used prospective cohort data [14]. However, the data were derived from MSM in the United States, and because of differences in the distribution of social and cultural demographics and risk factors among MSM, this model is not directly applicable to MSM in China [14].

Social media is an important communication tool to build virtual communities and networks, especially among the MSM community. WeChat is the most popular messaging and social media app in China; it had over 1 billion monthly active users in 2018 [15]. An online survey conducted among Chinese MSM reported that 57.9% of participants used online dating apps, which are associated with risky sexual behaviors and may foster a virtual environment that increases the risk of sexually transmitted diseases [16]. However, most published HIV risk prediction models for MSM are circulated through webpages and not through newer social media platforms.

Objectives

This study aims to build a prospective cohort-derived longitudinal HIV seroconversion risk assessment tool that is based on a popular social media platform to provide easy access for MSM to determine their personalized HIV incidence risk and inform targeted interventions.

Methods

Study Design and Participants

The open prospective cohort enrolled MSM from Shenyang, China, through a snowball sampling method and followed this cohort from January 2009 to January 2016 through the voluntary counseling and testing center in the First Affiliated Hospital of China Medical University [17]. The MSM cohort was recruited from baths, bars, a social media app for the gay community, and other venues as conducted by community-based organization leaders. Initial participants were asked to recruit partners or peers to participate in the survey and given 50 yuan (approximately US $8) for each recruited participant. Following written informed consent, eligible participants were interviewed face-to-face by a trained staff member in a private counseling room. The interview asked about demographics, sexual practices, and substance use, including the history of recreational drug use and whether the participant used poppers and/or methamphetamine in the past 3 months. Condoms and lubricants were freely distributed to each MSM participant. Both HIV-1 and syphilis tests were conducted on each participant. Pretest and posttest counseling were provided to each participant at both the baseline and follow-up HIV-1 and syphilis tests. The cohort inclusion criteria were the following: (1) aged 15 years or older, (2) male who self-reported anal and/or oral intercourse experience with a male partner in the past 6 months, (3) baseline negative HIV antibody and nucleic acid screening, and (4) informed consent signed by themselves or their guardian. The study participants were assigned a unique six-digit personal identification code, which was used to link their test results to demographic information.

Potential Risk Factors of HIV Incidence

Factors that may be associated with HIV seroconversion among MSM in this cohort include demographic characteristics, sexual practices, sexual role, condom use during anal intercourse, number of male sexual partners, and history of recreational drug use. We used a recall window of the past 3 months. We used the HIV/AIDS-related knowledge questionnaire, which consisted of eight questions, and 1 point was given for the correct answer to each question. The questionnaire was the following:

XSL•FO
RenderX

1. Is it possible for a healthy person to be infected with HIV?
2. Can blood or blood products with HIV virus spread HIV?
3. Can sharing needles with others transmit HIV infection?
4. Can condom usage reduce the risk of HIV infection?
5. Does having a monogamous HIV-negative partner decrease the risk of HIV acquisition?
6. Can HIV be transmitted from an HIV-positive pregnant woman to her baby?
7. Can HIV be transmitted through having dinner with an HIV-positive individual?
8. Can HIV be spread through mosquito bites? [18]

We defined the outcome of HIV seroconversion as a baseline HIV antibody seronegative case that seroconverted to an HIV antibody seropositive case during the follow-up period. We defined the HIV seroconversion time as the midpoint date between the last tested HIV-seronegative date and the first tested HIV-seropositive date.

Laboratory Testing

Both HIV and syphilis testing were performed every 3 months during follow-up. The HIV screening test was performed through enzyme-linked immunosorbent assay (ELISA), and suspected positive cases were further confirmed by Western blot. HIV-1 antibody-negative cases and positive cases in which Western blot was uncertain or negative were further confirmed negative through pooled real-time polymerase chain reaction (Cobas Amplicor HIV-1 MONITOR Tom Test, v1.5, Roche, 21118390123). Syphilis serology was performed using the rapid plasma reagin test (RPR; Shanghai Kehua, China), and positive cases were further confirmed by the Treponema pallidum particle assay (TPPA, Serodia, Japan). Participants with plasma positive for both RPR and TPPA were deemed to be currently infected with syphilis.

Statistical Analysis and WeChat Applet Construction

The overall dataset was randomly divided into training and validation datasets at the approximate ratio of 2:1. Variables were selected using the backward variable selection method based on the Akaike information criterion in the Cox regression model. Variables with a *P* value less than .25 in the univariable Cox regression model were entered into the multivariable regression for variable selection, and variables with a *P* value less than .05 were retained in the final model. We used a bootstrap resampling procedure of 10,000 samples to test the stability of the predictive score in the training and validation datasets. We also used the Cox.zph function in the "survival" package of R to test the validity of the proportional hazards assumption. Internal and external consistency of the discrimination and calibration performance measures were evaluated by the bootstrap resampling procedure. Discrimination was evaluated using Harrell's concordance statistic (C-index), and calibration was conducted by comparing the actual proportion of those without incident HIV infection with the predicted probability of no incident HIV infection developed from the Kaplan-Meier estimates for each decile. We used multiple imputation methods to address missing values. If the variable had more than 50% missing, it was discarded. Based on the predictive model with the identified risk factors, a nomogram of the prediction model was constructed to allow visual estimation of the predicted 2- and 4-year cumulative risk. Alternatively, the risk of HIV infection could be calculated using a WeChat-based HIV risk calculator. The WeChat-based HIV risk calculator was designed as follows: (1) the system adopts client-server structure to communicate with transmission control protocol and the Internet Protocol protocol; (2) the server adopts model-view-controller design and the client is developed by WeiXin Markup Language, JavaScript based on the WeChat platform; (3) build relational database MySQL for data storage; and (4) build and run an HIV risk assessment software program on a WeChat applet compatible with iOS and Android mobile devices. We developed three stratified risk subgroups using the tertial cutoff points of the linear prediction value for risk stratification [19]. The cumulative incidences of HIV seroconversion among the three risk subgroups were compared by the log-rank test. The predicted 2- and 4-year absolute risks were calculated from the baseline probability, and the relative risk profile was developed from the Cox proportional hazards regression model. All statistical analyses were performed using SAS 9.4 (Cary, NC, USA) and R software version 2.13.2. A two-sided *P* value of less than .05 was considered statistically significant. The study protocol was approved by the Institutional Review Board of the First Affiliated Hospital of China Medical University, Shenyang, China ([2011]-36).

Results

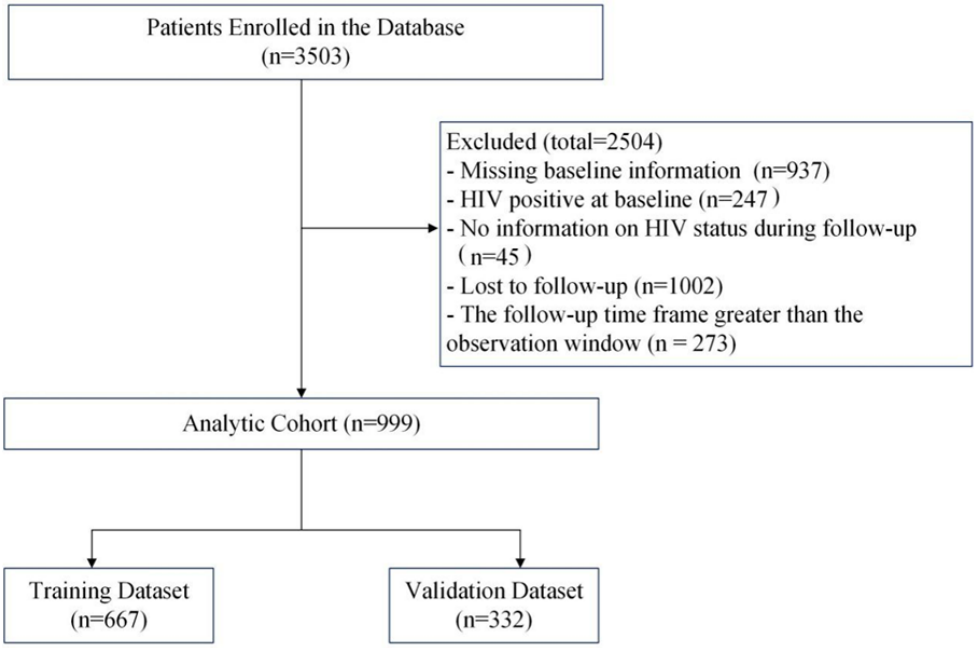
Selection and Characteristics of the Cohort

We examined 3503 medical records from a prospective open cohort of MSM followed from January 2009 to January 2016 in Shenyang, China. After excluding participants who had no follow-up data or were HIV positive at baseline, 999 MSM were included for the model construction. Of these, 667 MSM were randomly placed in the training dataset, and 332 were randomly placed in the validation dataset (Figure 1).

In the overall dataset, the mean age of the MSM was 27.5 (SD 9) years, and the range was 15 to 68 years. Overall, 48.7% (487/999) were aged 24 years or younger, 65.1% (650/999) had a local household registration, 85.8% (857/999) were ethnically Han, 65.3% (652/999) had a monthly income of less than US $430, 75.3% (752/999) were single, 39.8% (398/999) had a college and above education level, and 55.9% (558/999) had an AIDS knowledge score of less than 8 points. There were no statistically significant differences between the characteristics of the overall dataset, the training dataset, and the validation dataset, suggesting that the randomization into each subset worked well (Table 1).

XSL•FO
RenderX

Figure 1. Flowchart of study selection criteria from a prospective cohort of men who have sex with men from 2009 to 2016 in Shenyang, China.



XSL•FO
RenderX

Table 1. Characteristics of MSM participants in the overall, training, and validation dataset.

Characteristics and subgroup	Overall dataset (N=999), n (%)	Training dataset (n=667), n (%)	Validation dataset (n=332), n (%)	*P* value
Age (years)				.42
≤24	487 (48.7)	317 (47.5)	170 (51.2)	
>24	512 (51.3)	350 (52.5)	162 (48.8)	
Local residence				.67
Yes	650 (65.1)	431 (64.6)	219 (66.0)	
No	349 (34.9)	236 (35.4)	113 (34.0)	
Ethnicity				.82
Non-Han	142 (14.2)	96 (14.4)	46 (13.9)	
Han	857 (85.8)	571 (85.6)	286 (86.1)	
Monthly income (US$)				.54
<430	652 (65.3)	431 (64.6)	221 (66.6)	
≥430	347 (34.7)	236 (35.4)	111 (33.4)	
Education				.81
Less than high school	342 (34.2)	224 (33.6)	118 (35.5)	
High school	259 (25.9)	176 (26.4)	83 (25.0)	
College and above	398 (39.8)	267 (40.0)	131 (39.5)	
Marital status				.48
Single	752 (75.3)	507 (76.0)	245 (73.8)	
Married or cohabiting with a partner	247 (24.7)	160 (24.0)	87 (26.2)	
AIDS knowledge scores				.64
<8	558 (55.9)	376 (56.4)	182 (54.8)	
8	441 (44.1)	291 (43.6)	150 (45.2)	
Age of sexual debut with males (years)				.58
<30	917 (91.8)	610 (91.5)	307 (92.5)	
≥30	82 (8.2)	57 (8.5)	25 (7.5)	
Main venue used to seek male sexual partners				.32
Internet	527 (52.8)	360 (54.0)	167 (50.3)	
Bars/dance halls	43 (4.3)	31 (4.6)	12 (3.6)	
Parks/public baths	429 (42.9)	276 (41.4)	153 (46.1)	
Had regular male sexual partners				.59
Yes	571 (57.2)	385 (57.7)	186 (56.0)	
No	428 (42.8)	282 (42.3)	146 (44.0)	
Had casual male sexual partners				.74
Yes	597 (59.8)	396 (59.4)	201 (60.5)	
No	402 (40.2)	271 (40.6)	131 (39.5)	
Number of male sexual partners				.93
<3	612 (61.3)	408 (61.2)	204 (61.4)	
≥3	387 (38.7)	259 (38.8)	128 (38.6)	
Had condomless insertive anal intercourse				.52
Yes	384 (38.4)	261 (39.1)	123 (37.0)	
No	615 (61.6)	406 (60.9)	209 (63.0)	

XSL•FO
RenderX

Characteristics and subgroup	Overall dataset (N=999), n (%)	Training dataset (n=667), n (%)	Validation dataset (n=332), n (%)	*P* value
Had condomless receptive anal intercourse				.51
Yes	351 (35.1)	239 (35.8)	112 (33.7)	
No	648 (64.9)	428 (64.2)	220 (66.3)	
Rush poppers use				.92
Yes	100 (10.0)	65 (9.7)	35 (10.5)	
No	538 (53.9)	361 (54.1)	177 (53.3)	
Not available	361 (36.1)	241 (36.1)	120 (36.1)	
Tested positive for syphilis				.55
Yes	81 (8.1)	57 (8.5)	24 (7.2)	
No	685 (68.6)	450 (67.5)	235 (70.8)	
Not tested	233 (23.3)	160 (24.0)	73 (22.0)	

Cox Regression Analysis and Risk Score Formula

Table 2 lists the independent predictors for HIV seroconversion among MSM with the hazard ratios calculated by the multivariable Cox proportional hazards regression model. The hazard was lower for having condomless insertive anal intercourse (beta=−1.51, adjusted hazard ratio [aHR]=0.22, 95% CI 0.09-0.55, *P*=.001) compared to not having condomless insertive anal intercourse. The hazard was higher for having condomless receptive anal intercourse (beta=1.10, aHR=3.01, 95% CI 1.48-6.16, *P*=.003) compared to not having condomless receptive anal intercourse. The hazards were higher for either using the internet as the main venue to seek male sexual partners (beta=0.99, aHR=2.70, 95% CI 1.02-7.14, *P*=.046) or for using bars or dance halls as the main venue to seek male sexual partners (beta=2.06, aHR=7.84, 95% CI 2.64-22.91, *P*=.03) compared to using parks or public baths as the main venue to seek male sexual partners, and for rush poppers use (beta=0.88, aHR=2.40, 95% CI 1.10-5.27, *P*=.03) compared to not using rush poppers. The HIV prediction Cox regression model was thus the following:

$$F(t)=1-[S_0(t)]^{\exp[(-1.51)\times x1+1.10\times x2+0.99\times x3a+2.06\times x3b+0.88\times x4)]}$$

where $F(t)$ is the risk function or the probability of incident HIV infection over time (in years), $S_0(t)$ is the baseline survival function; *x1* is had condomless insertive anal intercourse (0=no, 1=yes), *x2* is had condomless receptive anal intercourse (0=no, 1=yes), *x3a* is the dummy variable that the main venue used to seek male sexual partners is the internet (1=internet, 0=parks/public baths), and *x3b* is the dummy variable that the main venue used to seek male sexual partners is bars or dance halls (1=bars/dance halls, 0=parks/public baths), and *x4* is rush poppers use (0=no, 1=yes). From the Kaplan-Meier estimation, the HIV-uninfected probabilities, S_0, were calculated as 0.96 for baseline, 0.90 for 1 year, 0.86 for 2 years, 0.80 for 3 years, and 0.74 for 4 years.

Model Discrimination and Calibration

The C-index was 0.75 (95% CI 0.65-0.85) in the training dataset, and 0.60 (95% CI 0.45-0.74) in the validation dataset. The calibration plot for the probability of no HIV infection at 2 years and at 4 years showed good agreement between predicted risk and the actual proportion of no HIV infection in both the training and validation datasets (Figure 2).

XSL•FO
RenderX

Table 2. Cox regression analysis of HIV seroconversion among men who have sex with men in the training dataset.

Characteristics and subgroup	Univariate analysis		Multivariate analysis	
	cHR[a] (95% CI)	*P* value	aHR[b] (95% CI)	*P* value
Age (years)				
≤24	1.00			
>24	0.98 (0.62-1.54)	.92	__[c]	—
Local residence				
Yes	1.00			
No	0.97 (0.61-1.54)	.89	—	—
Ethnicity				
Non-Han	1.00			
Han	1.09 (0.58-2.07)	.78	—	—
Monthly income (US$)				
<430	1.00			
≥430	0.89 (0.70-1.14)	.36	—	—
Education				
Less than high school	1.00			
High school	0.72 (0.40-1.31)	.29	—	—
College and above	1.09 (0.65-1.84)	.75	—	—
Marital status				
Single	1.00			
Married or cohabiting with a partner	0.50 (0.27-0.91)	.02	—	—
AIDS knowledge scores				
<8	1.00			
8	0.86 (0.53-1.39)	.54	—	—
Age of sexual debut (years)				
< 30	1.00			
≥ 30	0.14 (0.02-0.98)	.02	—	—
Main venue used to seek male sexual partners				
Internet	1.81 (1.08-3.04)	.02	2.70 (1.02-7.14)	.046
Bars/dance halls	5.58 (2.09-14.93)	<.001	7.84 (2.64-22.91)	.03
Parks/public baths	1.00			
Number of male sexual partners				
<3	1.00			
≥3	2.06 (1.29-3.28)	.002	—	—
Had condomless insertive anal intercourse				
Yes	0.66 (0.41-1.07)	.09	0.22 (0.09-0.55)	.001
No	1.00		1.00	
Had condomless receptive anal intercourse				
Yes	1.60 (1.00-2.50)	.047	3.01 (1.48-6.16)	.003
No	1.00		1.00	
Rush poppers use				
Yes	2.34 (1.10-5.00)	.004	2.40 (1.10-5.27)	.03

XSL•FO
RenderX

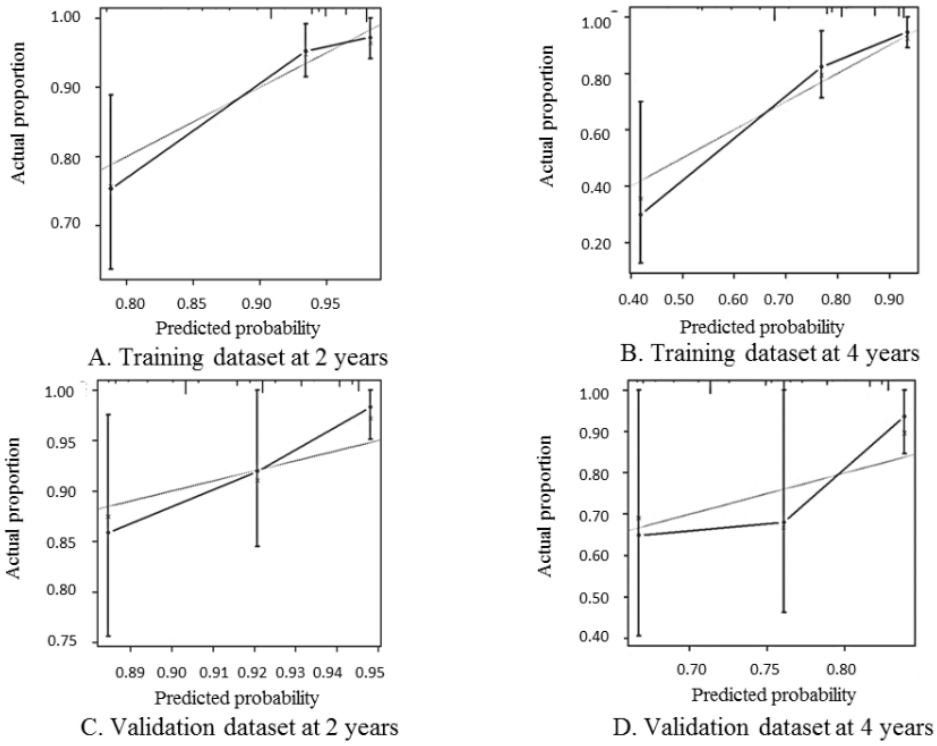
Characteristics and subgroup	Univariate analysis		Multivariate analysis	
	cHR[a] (95% CI)	*P* value	aHR[b] (95% CI)	*P* value
No	1.00		1.00	
Tested positive for syphilis				
Yes	1.59 (0.85-2.98)	.15	—	—
No	1.00			

[a]cHR: crude hazard ratio.

[b]aHR: adjusted hazard ratio.

[c]Not applicable.

Figure 2. Calibration plot of the HIV seroconversion risk prediction model for the training dataset at (A) 2 years and (B) 4 years, and for the validation dataset at (C) 2 years and (D) 4 years. The x-axes present the predicted probability of no incident HIV infection; the y-axes present the actual probability of no incident HIV infection. The gray lines are the actual proportion of men who have sex with men without incident HIV infection, the black lines are predicted probabilities, and the vertical capped lines are estimates with the 95% confidence intervals.



Development of the HIV Risk Assessment Tools

To facilitate the application of our model by public health workers and by the MSM community, we constructed a nomogram and a social media platform–based calculator. A nomogram is a visual risk assessment tool that allows for the approximate graphical computation of a mathematical function, and it integrates all the significant independent factors for HIV seroconversion determined by the training dataset to predict HIV seroconversion risk at multiple time points in the future (Figure 3). As WeChat is the most popular social media app in China, we also developed a WeChat-based risk calculator that is accessible through a QR code to facilitate distribution of our risk prediction model (Multimedia Appendix 1). Using the nomogram or the WeChat-based risk calculator, MSM individuals can enter their own HIV-related behavioral characteristics and quickly learn their objective risk of HIV infection in the future.

To facilitate the use of these two tools, tertile cutoff points of the linear predictor of risk based on the Cox regression were used to stratify MSM into low-, intermediate-, or high-risk subgroups, and the cutoff points were determined at 0.10 and 0.77. The log-rank test indicated a significant difference for the probability of no incident HIV infection among the three subgroups in the training dataset (P<.001), and a different trend among three subgroups in the validation dataset (P=.19; Figure 4).

XSL•FO
RenderX

Figure 3. Nomogram for incident HIV infection among men who have sex with men. To use the nomogram, the health official first determines and then sums the points of each variable located on the top point scale. Based on this sum, the health official uses the bottom point scale to determine the probability of HIV incidence. CIAI: condomless insertive anal intercourse; CRAI: condomless receptive anal intercourse.

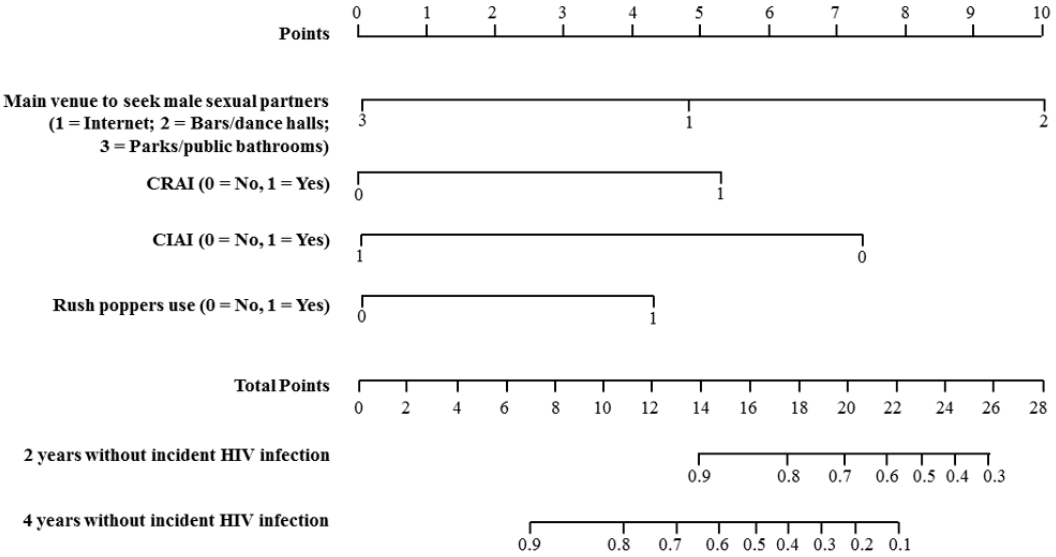
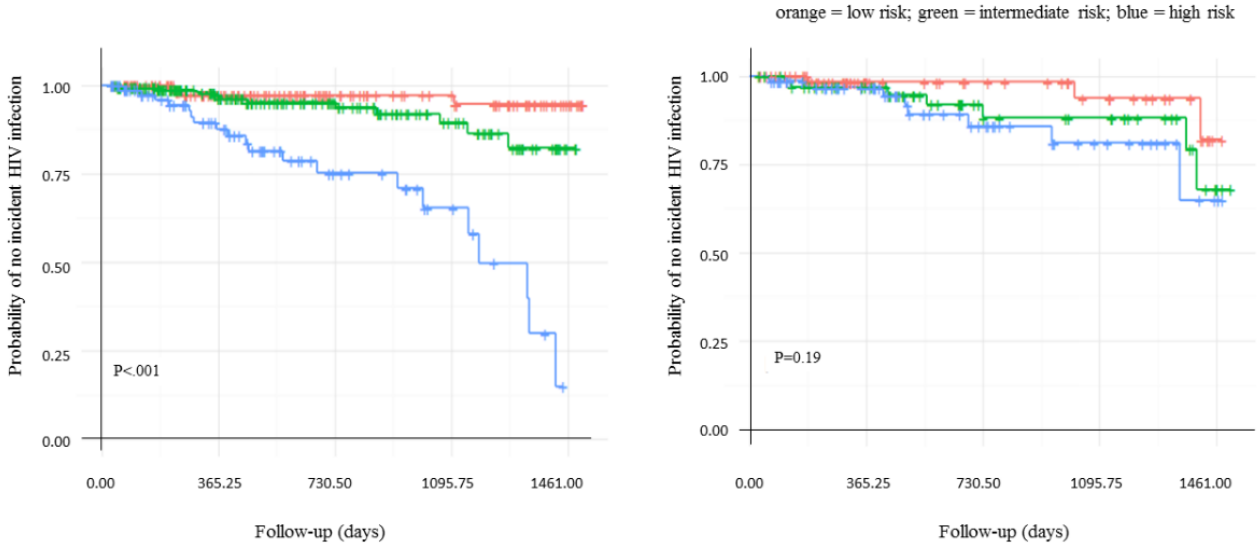


Figure 4. Risk stratification through Kaplan-Meier survival curves for (A) the training dataset and (B) the validation dataset.



Discussion

Principal Findings

This study constructed an HIV incidence risk assessment model based on data from an open prospective MSM cohort and developed two tools to convey this personalized HIV acquisition risk (ie, the nomogram and WeChat-based risk calculator). These tools can provide MSM their accurate objective risk of incident HIV infection over time and thus can stratify MSM into distinct risk subgroups (ie, low, intermediate, or high risk). The WeChat applet was built and distributed through the most popular social media platform in China to facilitate the use of this tool.

Our study found that rush poppers use is an independent predictor for HIV seroconversion among MSM. Studies have reported that rush poppers use was associated with sexual behaviors that have a high risk of HIV infection, such as multiple sexual partners, commercial sex, and group sex [20]. Thus, our inclusion of this risk factor for HIV seroconversion in our model is concurrent with previous findings. Similar to the Menza model which used Seattle & King County STD clinic electronic records to estimate hazard ratios for HIV acquisition

XSL•FO
RenderX

based on Cox proportional hazards models, unprotected anal intercourse with a partner of positive or unknown HIV status as the acknowledged predictors of HIV seroconversion in Menza score [14], we also find this predictor to be significant in our multivariable analysis. Therefore, partners of individuals found to be HIV positive should be notified promptly by the health department and offered testing and treatment to diagnose and reduce secondary HIV transmission [21].

We also found that compared to the main venue for seeking male sexual partners being parks or public baths, the main venue being bars or clubs or being the internet were independent predictors of HIV seroconversion among MSM. The Guangdong HIV prediction model found that MSM among different venues (ie, parks, public baths, bars/clubs, or internet) had different high-risk sexual behaviors and HIV infection risks [12]. For example, since sex workers often recruit customers at bars or clubs, MSM who use bars or clubs as their main venue to seek male sexual partners may have more commercial sex [22]. In addition, MSM who use bars or clubs as their main venue to seek male sexual partners are younger compared to MSM who use parks or public baths [23]. Therefore, to efficiently address the HIV epidemic among young MSM, HIV prevention strategies must be implemented in bars or clubs. Moreover, since male sexual partners are more readily accessible because of the internet and social media facilitating connections, MSM who use the internet to seek male sexual partners have more sexual partners and unprotected anal intercourse [24-28] than MSM who use other venues to find male sexual partners. Thus, researchers should take advantage of the internet and social media to advertise and distribute tailored comprehensive intervention packages.

Previous models used odds ratios as the scoring value; whereas, we used hazard ratios. Compared to odds ratios, which are cumulative over a time span with a defined endpoint, hazard ratios represent the instantaneous risk over the study period, and thus are easier to understand and interpret by clinicians and the MSM community. In addition, our tool, in contrast to previously published risk assessment tools, can predict the incident HIV infection risk in the next 1, 2, 3, and 4 years.

Although the discriminatory accuracy of our model was modest, the C-index estimates that we report are slightly higher than those of other risk models, such as the Menza score [14], SDET score [11], which were also commonly used to guide clinical decisions. This cumulative risk evaluation over time accurately provides MSM the impact of their high-risk sexual behaviors on their personal HIV infection risk. Finally, our tools can be applied in other settings and countries based on our methods. Other researchers could use their local MSM prospective cohort data to calculate their own parameters to develop their specialized HIV risk assessment tool. In addition, researchers in countries with a large social media presence should also consider advertising and distributing these tools through popular social media platforms.

We constructed a personalized and objective model, and further developed a nomogram, a graphical calculating device widely used in tumor prediction models [29,30], and a WeChat-based HIV risk assessment, which can be used on mobile phones or tablets to facilitate the distribution and application of our model to the MSM community. We designed our tools to be easily accessible and distributed through a social media platform compared with traditional internet-based HIV prediction tools [31]. We based this tool on the most popular social media platform in China, WeChat, because MSM frequently use this platform to communicate, seek sexual partners, and share information and thus can reach all members of the MSM social network [32,33]. In addition, the WeChat applet is compatible with iOS and Android operating systems, and our risk prediction tool can be easily combined with the existing WeChat functions facilitating online and offline HIV prevention and treatment. Our WeChat risk prediction app is readily accessible and the MSM community can easily determine their personalized HIV infection risk and be connected to comprehensive health interventions [34]. Since 2017, we recorded 4158 visits to our WeChat applet. We are currently conducting a randomized controlled clinical trial to assess the effectiveness of this HIV risk prediction tool at promoting HIV testing and decreasing HIV-related high-risk behaviors, such as reducing the number of sexual partners, increasing the proportion of condom usage, etc. This online comprehensive intervention based on the HIV risk prediction model could reduce the number of homosexual partners and promote the use of condoms with casual partners in the MSM population.

Limitations

Our WeChat tool can only be used through WeChat, and thus is only available in China and neighboring countries. However, as our model is based on a cohort of Chinese MSM, we expect our tool to be most accurate when applied in China. Second, the current HIV infection risk assessment software gives related prevention and control recommendations but is not connected to more substantive HIV interventions. Future steps include integrating this risk assessment tool with other ongoing HIV detection and HIV high-risk behavior intervention projects for MSM to improve risk perception and promote regular HIV testing and HIV prevention services. We also used snowball sampling and thus may have sampling bias as the social network collected is not random and may be limited to a specific group or geographic area. Finally, the recruitment time period spans 8 years and characteristics of the MSM community and risk factors of HIV infection may have changed over this time period; therefore, our results may be less accurate than a larger cohort followed for a shorter period of time.

Conclusions

We developed and validated an HIV incidence risk assessment model for MSM in China. This model provides objective self-assessments and predictions of incident HIV infection risk. We then developed this model into two separate tools with one tool built on the most popular social media platform in China. MSM can use these tools to quantify their personal HIV infection risk. In addition, public health officials and community health workers can use these tools to conduct accurate quantitative HIV risk assessments of their MSM patients and thus determine HIV high-risk participants to target for risk reduction interventions in hopes of mitigating the HIV epidemic among MSM.

XSL•FO
RenderX

Acknowledgments

We would like to thank all the participants in this cohort of Chinese MSM from Shenyang, China, for their participation. This work was supported by the Mega-Projects of National Science Research (13th Five-Year Plan, 2017ZX10201101-002-007), National Natural Science Foundation of China (81872674), National Science and Technology Major Project (2018ZX10101-001-001-003), and the Central Public Interest Scientific Institution Basal Research Fund of Chinese Academy of Medical Sciences (2018PT31042).

Authors' Contributions

KY and JJX contributed equally to this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Risk prediction WeChat applet instructions and architecture.

[PPTX File, 133KB - jmir_v21i6e13475_app1.pptx]

References

1. UNAIDS Joint United Nations Programme on HIV/AIDS. The Gap Report. Geneva: UNAIDS; 2014. URL: http://files.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf [accessed 2019-01-01] [WebCite Cache ID 75d8VkUvg]
2. Zhang W, Xu J, Zou H, Zhang J, Wang N, Shang H. HIV incidence and associated risk factors in men who have sex with men in Mainland China: an updated systematic review and meta-analysis. Sex Health 2016 May 26:373-382. [doi: 10.1071/SH16001] [Medline: 27294689]
3. UNAIDS Joint United Nations Programme on HIV/AIDS. 90-90-90: An Ambitious Treatment Target to Help End the AIDS Epidemic. Geneva: UNAIDS URL: http://www.unaids.org/sites/default/files/media [accessed 2018-10-30] [WebCite Cache ID 75dARUtdd]
4. Ministry of Health of the People's Republic of China. 2012 China AIDS Response Progress Report. 2012 Mar 31. URL: http://www.unaids.org/sites/default/files/country/documents//file,68497,es [accessed 2019-05-22] [WebCite Cache ID 75dAYlosl]
5. Han L, Bien CH, Wei C, Muessig KE, Yang M, Liu F, et al. HIV self-testing among online MSM in China: implications for expanding HIV testing among key populations. J Acquir Immune Defic Syndr 2014 Oct 01;67(2):216-221 [FREE Full text] [doi: 10.1097/QAI.0000000000000278] [Medline: 24991972]
6. MacKellar DA, Valleroy LA, Secura GM, Behel S, Bingham T, Celentano DD, Young Men's Survey Study Group. Perceptions of lifetime risk and actual risk for acquiring HIV among young men who have sex with men. AIDS Behav 2007 Mar;11(2):263-270. [doi: 10.1007/s10461-006-9136-0] [Medline: 16791527]
7. Molina J, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, et al. On-demand preexposure prophylaxis in men at high risk for HIV-1 infection. N Engl J Med 2015 Dec 3;373(23):2237-2246. [doi: 10.1056/NEJMoa1506273] [Medline: 26624850]
8. Li L, Kuang C, Qi J. Survey of HIV infection status and risk factors in the MSM population in Luohu District, Shenzhen City. China Public Health Management 2014;30(3):384-386. [doi: 10.19568/j.cnki.23-1318.2014.03.031]
9. Haukoos J, Lyons M, Lindsell C, Hopkins E, Bender B, Rothman RE, et al. Derivation and validation of the Denver Human Immunodeficiency Virus (HIV) risk score for targeted HIV screening. Am J Epidemiol 2012 Apr 15;175(8):838-846 [FREE Full text] [doi: 10.1093/aje/kwr389] [Medline: 22431561]
10. Wahome E, Fegan G, Okuku HS, Mugo P, Price MA, Mwashigadi G, et al. Evaluation of an empiric risk screening score to identify acute and early HIV-1 infection among MSM in Coastal Kenya. AIDS 2013;27(13):2163-2166. [doi: 10.1097/QAD.0b013e3283629095]
11. Hoenigl M, Weibel N, Mehta SR, Anderson CM, Jenks J, Green N, et al. Development and validation of the San Diego Early Test score to predict acute and early HIV infection risk in men who have sex with men. Clin Infect Dis 2015 Apr 22;61(3):468-475. [doi: 10.1093/cid/civ335]
12. Hu P, Zhong F, Cheng W, Xu HF, Ling L. [Study on the infectious risk model of AIDS among men who have sex with men in Guangzhou]. Zhonghua Liu Xing Bing Xue Za Zhi 2012 Jul;33(7):667-671. [Medline: 22968013]
13. Yin L, Zhao Y, Peratikos MB, Song L, Zhang X, Xin R, et al. Risk prediction score for HIV infection: development and internal validation with cross-sectional data from men who have sex with men in China. AIDS Behav 2018 May 21;22(7):2267-2276. [doi: 10.1007/s10461-018-2129-y]

XSL•FO
RenderX

14. Menza TW, Hughes JP, Celum CL, Golden MR. Prediction of HIV acquisition among men who have sex with men. Sex Transm Dis 2009 Sep;36(9):547-555 [FREE Full text] [doi: 10.1097/OLQ.0b013e3181a9cc41] [Medline: 19707108]
15. Zeng F, Deng G, Wang Z, Liu L. WeChat: a new clinical teaching tool for problem-based learning. Int J Med Educ 2016 Apr 25;7:119-121 [FREE Full text] [doi: 10.5116/ijme.5708.e5c4] [Medline: 27111920]
16. Tang W, Best J, Zhang Y, Liu F, Tso LS, Huang S, et al. Gay mobile apps and the evolving virtual risk environment: a cross-sectional online survey among men who have sex with men in China. Sex Transm Infect 2016 Jun 09;92(7):508-514. [doi: 10.1136/sextrans-2015-052469]
17. Xu J, Yu H, Tang W, Leuba SI, Zhang J, Mao X, et al. The effect of using geosocial networking apps on the HIV incidence rate among men who have sex with Men: eighteen-month prospective cohort study in Shenyang, China. J Med Internet Res 2018 Dec 21;20(12):e11303 [FREE Full text] [doi: 10.2196/11303] [Medline: 30578225]
18. Yang X, Attané I, Li S, Yang B. Same-sex sexual behaviors among male migrants in a context of male. Am J Mens Health 2012 Nov;6(6):485-496. [doi: 10.1177/1557988312453479] [Medline: 22782362]
19. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. In: Statistics for Biology and Health. New York: Springer; 2009.
20. Buchbinder SP, Vittinghoff E, Heagerty PJ, Celum CL, Seage GR, Judson FN, et al. Sexual risk, nitrite inhalant use, and lack of circumcision associated with HIV seroconversion in men who have sex with men in the United States. J Acquir Immune Defic Syndr 2005 May 01;39(1):82-89. [Medline: 15851918]
21. Ko N, Tseng P, Huang Y, Chen Y, Hsu S. Seeking sex partners through the internet and mobile phone applications among men who have sex with men in Taiwan. AIDS Care 2016 Dec;28(7):927-931. [doi: 10.1080/09540121.2015.1131969] [Medline: 26754350]
22. Gao Y, Liu Y, Li S. Prevalence and predictors of HIV, syphilis and herpes simplex type 2 virus (HSV-2) infections among the men who have sex with men(MSM) in Beijing. Chinese J Public Health 2012;28(4):3. [doi: 10.11847/zgggws2012-28-04-15]
23. Grov C. HIV risk and substance use in men who have sex with men surveyed in bathhouses, bars/clubs, and on Craigslist.org: venue of recruitment matters. AIDS Behav 2012 May;16(4):807-817 [FREE Full text] [doi: 10.1007/s10461-011-9999-6] [Medline: 21748276]
24. Shi L, Wei C, McFarland W, Yan H, Li J, Raymond HF. Comparing samples of men who have sex with men recruited online and in venues, Jiangsu Province, China, 2013. LGBT Health 2016 Jun;3(3):238-242 [FREE Full text] [doi: 10.1089/lgbt.2014.0129] [Medline: 26651841]
25. Bull S, McFarlane M. Soliciting sex on the Internet: what are the risks for sexually transmitted diseases and HIV? Sex Transm Dis 2000 Oct;27(9):545-550. [Medline: 11034529]
26. Ashton M, Sopwith W, Clark P, McKelvey D, Lighton L, Mandal D. An outbreak no longer: factors contributing to the return of syphilis in Greater Manchester. Sex Transm Infect 2003 Aug;79(4):291-293 [FREE Full text] [doi: 10.1136/sti.79.4.291] [Medline: 12902577]
27. Klausner J, Wolf W, Fischer-Ponce L, Zolt I, Katz MH. Tracing a syphilis outbreak through cyberspace. JAMA 2000 Jul 26;284(4):447-449. [Medline: 10904507]
28. Tashima KT, Alt EN, Harwell JI, Fiebich-Perez DK, Flanigan TP. Internet sex-seeking leads to acute HIV infection: a report of two cases. Int J STD AIDS 2003;14(4):285-286. [doi: 10.1258/095646203321264926]
29. Huang G, Dong Z, Xie X. Establishment of a nomogram for an individualized predictive model of osteosarcoma prognosis. Chinese J Orthopaed 2015;35(2):133-141. [doi: 10.3760/cma.j.issn.0253-2352.2015.02.006]
30. Tang L, Li C, Li J, Chen W, Chen Q, Yuan L, et al. Establishment and validation of prognostic nomograms for endemic nasopharyngeal carcinoma. J Natl Cancer Inst 2016 Jan;108(1):1. [doi: 10.1093/jnci/djv291] [Medline: 26467665]
31. Hanna R, Rohm A, Crittenden VL. We're all connected: the power of the social media ecosystem. Bus Horizons 2011 May;54(3):265-273. [doi: 10.1016/j.bushor.2011.01.007]
32. Cao B, Liu C, Durvasula M, Tang W, Pan S, Saffer AJ, et al. Social media engagement and HIV testing among men who have sex with men in China: a nationwide cross-sectional survey. J Med Internet Res 2017 Jul 19;19(7):e251. [doi: 10.2196/jmir.7251]
33. Weebly. Advantages of WeChat URL: https://s222248.weebly.com/advantage?disadvantage.html [accessed 2019-05-22] [WebCite Cache ID 78YfHYHu7]
34. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth and Uhealth 2015;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]

Abbreviations

aHR: adjusted hazard ratio
cHR: crude hazard ratio
ELISA: enzyme-linked immunosorbent assay
MSM: men who have sex with men

XSL•FO
RenderX

PrEP: preexposure prophylaxis
RPR: rapid plasma reagin test
TPPA: Treponema pallidum particle agglutination assay
SDET: San Diego Early Test

Edited by G Eysenbach; submitted 28.01.19; peer-reviewed by TH Kwan, MS Aslam, B Hou, F Pourmalek; comments to author 23.03.19; revised version received 25.04.19; accepted 02.05.19; published 18.06.19.

Please cite as:
Yun K, Xu J, Leuba S, Zhu Y, Zhang J, Chu Z, Geng W, Jiang Y, Shang H
Development and Validation of a Personalized Social Media Platform–Based HIV Incidence Risk Assessment Tool for Men Who Have Sex With Men in China
J Med Internet Res 2019;21(6):e13475
URL: https://www.jmir.org/2019/6/e13475/
doi:10.2196/13475
PMID:31215509

©Ke Yun, Junjie Xu, Sequoia Leuba, Yunyu Zhu, Jing Zhang, Zhenxing Chu, Wenqing Geng, Yongjun Jiang, Hong Shang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 18.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Examining Cost Measurements in Production and Delivery of Three Case Studies Using E-Learning for Applied Health Sciences: Cross-Case Synthesis

Edward Meinert[1,2], MA, MSc, MBA, MPA, PhD, CEng FBCS, EUR ING; Abrar Alturkistani[1], BA, MPH; Kimberley A Foley[1], BSc, BEd, MSc, PhD; David Brindley[2], MEng, MSc, DPhil; Josip Car[1], MD, PhD, FRCPE, FFPH

[1]Department of Primary Care and Public Health, Digital Global Health Unit, Imperial College London, London, United Kingdom

[2]Department of Paediatrics, Healthcare Translation Research Group, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Edward Meinert, MA, MSc, MBA, MPA, PhD, CEng FBCS, EUR ING
Healthcare Translation Research Group
University of Oxford
Department of Paediatrics
Children's Hospital
John Radcliffe Hospital
Oxford, OX3 9DU
United Kingdom
Phone: 44 7824446808
Email: edward.meinert@paediatrics.ox.ac.uk

Abstract

Background: The World Health Report (2006) by the World Health Organization conveys that a significant increase is needed in global health care resourcing to meet the current and future demand for health professionals. Electronic learning (e-Learning) presents a possible opportunity to change and optimize training by providing a scalable means for instruction, thus reducing the costs for training health professionals and providing patient education. Research literature often suggests that a benefit of e-Learning is its cost-effectiveness compared with face-to-face instruction, yet there is limited evidence with respect to the comparison of design and production costs with other forms of instruction or the establishment of standards pertaining to budgeting for these costs.

Objective: To determine the potential cost favorability of e-Learning in contrast to other forms of learning, there must first be an understanding of the components and elements for building an e-Learning course. Without first taking this step, studies lack the essential financial accounting rigor for course planning and have an inconsistent basis for comparison. This study aimed to (1) establish standard ingredients for the cost of e-Learning course production and (2) determine the variance instructional design has on the production costs of e-Learning courses.

Methods: This study made use of a cross-case method among 3 case studies using mixed methods, including horizontal budget variance calculation and qualitative interpretation of responses from course designers for budget variance using total quality management themes. The different implementation-specific aspects of these cases were used to establish common principles in the composition of budgets in the production and delivery of an applied health professional e-Learning course.

Results: A total of 2 case studies reported significant negative budget variances caused by issues surrounding underreporting of personnel costs, inaccurate resource task estimation, lack of contingency planning, challenges in third-party resource management, and the need to update health-related materials that became outdated during course production. The third study reported a positive budget variance because of the cost efficiency derived from previous implementation, the strong working relationship of the course project team, and the use of iterative project management methods.

Conclusions: This research suggests that the delivery costs of an e-Learning course could be underestimated or underreported and identifies factors that could be used to better control budgets. Through consistent management of factors affecting the cost of course production, further research could be undertaken using standard economic evaluation methods to evaluate the advantages of using e-Learning.

(J Med Internet Res 2019;21(6):e13574) doi:10.2196/13574

XSL•FO
RenderX

KEYWORDS

education; distance education; professional education; online education; online learning; costs and cost analysis; economics

Introduction

Rationale

The World Health Report (2006) by the World Health Organization (WHO) [1] conveys that a significant increase is needed in global health care resourcing to meet the current and future demand for health professionals. Current challenges to health care resourcing include the increasing demand resulting from the aging population's need for chronic disease management, in addition to the growing population placing an increased demand on primary care [2]. This increased demand on resources requires a scalable means to train resources; opportunities to optimize training through alternatives to face-to-face instruction present the possibility of increasing the pace and breadth of education to health care resourcing. A 2015 WHO systematic review of e-Learning for undergraduate health professional education concluded that "computer-based and Web-based e-Learning is no better and no worse than face-to-face learning with regards to knowledge and skill acquisition" [3]. e-Learning is defined as "an approach to teaching and learning, representing all or part of the educational model applied, that is based on the use of electronic media and devices as tools for improving access to training, communication, and interaction and that facilitates the adoption of new ways of understanding and developing learning" [4]. It presents a possible opportunity to change and optimize training in health professions (including clinical, allied, and applied health sciences, as well as patient education) by providing a scalable means for instruction, thus reducing the costs necessary in delivery and implementation. If we accept that pedagogically e-Learning can result in a positive educational effect used under optimal circumstances, which is still subject to ongoing investigation, there remains the possibility that deployment of e-Learning could affect the scale, cost, and reach of health professions education.

Research Problem

One of the motivations for implementing e-Learning is the potential long-term efficiency gain in its delivery model [5,6]. A course delivered digitally versus the cost of a lecturer providing face-to-face instruction appears to have long-term cost favorability [7]. The literature often suggests that a benefit of Web-based learning is its cost-effectiveness compared with face-to-face instruction [8]; however, there is limited evidence validating comparison with other forms of instruction or standards for the budgeting of the costs in the production and execution of e-Learning courses. In the case of massive open online courses (MOOCs), there is limited evidence on the costs associated with their production [9]. In addition, the costs to develop an e-Learning course are significant when executed to a high standard. Although there are studies that capture data relating to factors associated with educational costs, measurement in these studies are collected inconsistently and include a wide variety of factors [3,10]. There is limited transparency in costing models because of sensitivity on where direct costs should be applied [11]. A systematic means is required to comprehensively record costs that can then subsequently enable testing of whether the e-Learning course has desirable economic properties and under what scenarios [12]. If proven so, this could assist in addressing the high cost of delivering health professions education. By contrast, should evidence point the other way, having discrete data points will allow those involved in online health education to identify ways to optimize costs in delivery. The primary issue here is identification of the direct and indirect costs in implementation, which then allows the execution of further economic evaluation.

Aims and Objectives

This aim of this study was to establish an approach for identifying costs in the design, development, and deployment of applied health (defined as applied health subjects) sciences e-Learning courses and to subsequently propose a budgeting framework for the planning and management of e-Learning course implementations. The costs in this study include the direct and indirect costs from inception through course delivery. This approach will allow course designers and implementers to leverage knowledge gained from the study's e-Learning case studies across different implementation contexts to better plan and manage future implementations, which will also create a reusable framework to apply cost planning. This work will demonstrate the effect in pre-implementation budget management against the proposed framework and should result in better course planning.

The study's objectives are as follows:

- Establish an approach to capture standard components or ingredients for the cost of the production of an e-Learning course.
- Determine the effect that instructional design has on the production costs of e-Learning courses.

The study's aims and objectives intend to address a gap in the research literature concerning implementation details on planning and executing e-Learning in health professions education [8]. In addition to limited cost-centered studies on e-Learning for health professions education, there are limited details on how course designers and producers are calculating the associated costs for production of these course types. Developing models will allow for the adoption of data sharing and course planning for improved management in execution of this course method and for further refinement and analysis. To explore this issue, this research examines the following 3 distinct e- Learning implementations as case studies.

Educating Administrative Staff to Engage With Young Patients

The course was created as a small private online course (SPOC) to prepare general practice administrative staff for issues in the management of adolescents. The course used case studies to provide training to help general practice staff feel confident in helping adolescents with a goal of improving the patient experience.

XSL•FO
RenderX

The Impact of Climate Change on Public Health

This course was created as an MOOC to educate citizens on the relationship between climate change and public health by using a multidisciplinary academic framework in data science to analyze, interpret, and present evidence. Core case studies focused on climate change and its health economic effect on local, regional, and national health systems.

Data Science in Health Care Using Real-World Evidence

This course was created as a blended MOOC to make learners aware of the effect data science can have on medicine and inspire the application of these methods across various undergraduate curriculum disciplines, the UK National Health Service commissioning support organizations, health care regulation organizations, and life sciences industries (ie, pharmaceuticals, biotechnology, and medical devices).The implementation of the blended MOOC was executed as a face-to-face course for learners; learners first took part in the MOOC and were then offered a residential course examining case studies. The target audience of the MOOC was allied health professionals or citizens looking to transition or enhance skills in data science in health care–related industries such as the pharmaceutical industry or biotech organizations. One of the key objectives of the course was to establish a global network of people to continue and advance the dialogue on data science in health care. Some of the course outcomes include the use and application of real-world evidence data collection and analysis techniques in health care settings.

Methods

A mixed-methods case study design was selected to support a systematic means of observing the subject of investigation [13] and the ability to combine quantitative and qualitative approaches [14]. Mixed-methods research presents an opportunity to combine the strengths of quantitative and qualitative research to counteract the limitations inherent when each method is used in isolation [14]. In this study, for example, the limitations of quantitatively isolating cost differences in the 3 cases are strengthened by the repeatable and generalizable nature of the qualitative approach used to interpret results. Case studies were selected based on their relevance to the study inquiry and the ability to capture, record, and analyze data from each case. Each study was structured through a study protocol to govern the case execution.

Case Study Overview

Case Overview

The objective of the case study is to inform the way future costs are budgeted in the development of e-Learning courses. The research forms part of a broader investigation into the costs associated with e-Learning course production; the main focus of each case was to collect primary evidence in the construction of these costs to allow for further research comparing results with other Web-based learning implementation types.

- Study question: how are the total costs for the production and delivery of an e-Learning course (dependent on type) calculated?
- Proposition: actual and budgeted costs will vary in the production or delivery of this course type.

Existing research literature indicates challenges in the capture of total costs for the production of Web-based learning despite standard methods for cost calculation [8]. The reason for this variance is likely because the skills required to create instructional learning design and to capture costs are different, and educators are not trained in cost accounting methods.

The analytical framework for this investigation is based on the cost analysis methods underpinning education economic evaluation developed by Levin [15], which extends the standard costing and variance calculation principles of activity-based costing [16-18]. The *ingredients method* [15] is used to capture total cost production against cost categories. It examines the core composition of costs in the delivery of an education intervention; this is an activity-based costing approach that seeks to understand the core components required for delivery. Defining core costs is critical to performing further economic evaluations, though it is important to note that the scope of this research is limited to cost identification and not further economic analysis (eg, cost-benefit analysis, cost-effectiveness analysis, cost-utility analysis, and cost-feasibility analysis).

Case study protocols (Multimedia Appendices 1,2, and 3,) were developed at study commencement to demonstrate the way costs would be captured and analyzed. These protocols, in addition to a protocol for qualitative and quantitative analysis of learning effect (which is outside the scope of the cost investigation) [19] were drafted, submitted, peer reviewed, and approved.

Data Collection Procedures

Evidence to Be Expected

To validate the costs reported in the actual budget (which was an actual cost report), at least 2 separate sources confirming the final reported amount were sought (eg, for a reported incurred cost for staff, timesheets were reviewed to match hours to costs, task completion, and assignment in a project plan). These data comparisons increased the likelihood that reported data were accurate.

Events to Be Observed

Although the course implementation was observed and additional studies completed investigating the education effect, the scope of this study was centered on the cost decision making, and the way production affected cost delivery. Therefore, the observation scope for this study focused on reported costs and the way these correlated data to time actuals.

Documentation to Be Reviewed

The project budget, actual costs, and timesheets were reviewed for this study. Although there will be a review of the completed course and observation of the way the course uptake is completed, the latter shall be excluded from this study. A traceability log was maintained in Microsoft Excel linking the research questions to data sources and the study findings.

XSL•FO
RenderX

Protocol Questions

Study question: how are the total costs for the production and delivery of an e-Learning (type dependent on implementation type) course calculated?

- The costs will be measured and ingredients captured and analyzed to understand the factors affecting course production.
- Data will be collected to support the cost analysis categories.
- The corresponding evidence will be used to summarize ways that cost capture practices could be improved.

Study Framework

Plan

Each case study followed a 6-stage process in the investigation (Table 1) [13]. The research question centered on identifying the total costs of production and delivery in these e-Learning implementations, and the effect of factors on variance from anticipated budgets. It was selected because evidence from the literature suggests inconsistency in the determination of costs for the delivery of Web-based courses [20]. This is significant because the lack of consistent cost capture mechanisms for Web-based learning compromises any further evaluation. Despite available methods to avoid this outcome, the literature presents research with claims that Web-based learning is more *cost effective* than face-to-face learning. This research provides a structured means to generate evidence to subsequently evaluate such claims by collecting baseline data on course production for further evaluation.

Design

The research design (Table 2) was structured on 4 components (proposition, the case definition, logic linking data to the proposition, and criteria for interpreting findings) to explore the following research question: how are the total costs for the production and delivery of e-Learning calculated (with the e-Learning implementation type variant depending on the case study)? Given the inconsistency in the presentation of costs in the literature and recognizing that using budgets to determine the cost of educational delivery is insufficient [21], the governing proposition of the investigation was that there would be variance between the budgeted costs and the actual costs to produce the course. This was explored through cases that would examine the cost and the measurement of costs and place value on ingredients. Levin developed this *ingredients method* to capture and analyze the costs in the delivery of an educational program. To link the case to the proposition, the cost calculation was completed and then interpreted via a variance calculation of actual to budgeted costs, and rationales were developed to justify variations.

Examination of these cases provides data to analyze the relationship between course production and budgeting in the delivery of e-Learning and provides evidence for constructing accurate budget models.

Each case was tested for construct validity (testing that data sources come from multiple sources), external validity (testing that demonstrates how principal findings could be extensible) and reliability (testing that shows how the activities of the study can be replicated) to ensure data triangulation, the ability for study replication, and standardization for project data collection [13]. Ethical approval for each study was obtained through the Imperial College Education Ethics Research Committee (case 1: EERP1516-005; case 2 and 3: EERP1617-030).

Prepare

The investigation was focused on cost measurement and analysis, structured by 3 cost categories, and further subdivided using a 7-step process (illustrated in Table 3 below) to analyze the pre- and postproduction budget [21]. Levin's model uses an activity-based standard-costing accountancy approach, which assigns costs as they are consumed per implementation area [25,26].

Table 1. Case study framework.

Stage	Outcome
Plan	Case description and linking of case approach to investigation outcomes.
Design	Construction of research design and linkage of research questions, data, and criteria for evaluation and synthesis.
Prepare	Draft, execution, and approval of study protocols.
Collect	Data collection strategy executed from a *realist* perspective to capture the decision making of the course designers centered on cost attributes.
Analyze	Data extracted into categories for review and analyzed for variance calculation. Data analysis centers on 3 cost categories in the design of the preproduction budget submitted to the funder for each case. Category A: concept and measurement of costs: The preproduction budget was analyzed for the following ingredient categories: (1) personnel, (2) estate charges, (3) equipment and materials, (4) indirect costs, and (5) stakeholder costs; Category B: placing values on ingredients: With the full cost of production defined, values were associated with each ingredient subcategory to reflect the chargeable cost; Category C: calculating costs: To record a variance calculation, a comparison of the budget with the incurred costs was reviewed on a quarterly basis. Variance=Actual spending–Budgeted spending.
Share	The findings of the variance calculation and synthesis of analysis of reasons leading to variation were presented in a report for publication in a peer-reviewed journal. (This study).

XSL•FO
RenderX

Table 2. Case study research design.

Case (year)	Study question	Proposition	The case (definition)	Logic linking data to the proposition	Criteria for interpreting findings
Case 1: Educating administrative staff to engage with young patients (2016) [22]	How are the total costs for the production and delivery of this e-Learning course calculated?	Actual and budgeted costs will vary in the production/delivery of this course type	Determination and measurement of costs	Cost analysis of project, actual, and underreported costs	Variance calculation from the project budget
Case 2: The impact of climate change on public health (2017) [23]	How are the total costs for the production and delivery of this e-Learning course calculated?	Actual and budgeted costs will vary in the production/delivery of this course type	Determination and measurement of costs	Cost analysis of project, actual, and underreported costs	Variance calculation from the project budget
Case 3: Data science in healthcare using real world evidence (2018) [24]	How are the total costs for the production and delivery of this e-Learning course calculated?	Actual and budgeted costs will vary in the production/delivery of this course type	Determination and measurement of costs	Cost analysis of project, actual, and underreported costs	Variance calculation from the project budget

Table 3. Course production ingredients cost analysis.

Cost categories	Objectives—adapted from Levin (2001, 2018) [27,21]
Category A: concept and measurement of costs	Steps 1 to 5: Describe the concept of costs; show the inadequacy of budgets for cost analysis; present a methodology for measuring costs; identify categories of cost ingredients; describe sources of cost information
Category B: placing values on ingredients	Steps 6 and 7: Describe the purpose and principles for determining the values of ingredients; present methods for placing values on specific types of ingredients

Collect

Evidence from the course was retrieved from project documents and records of finance activity. The data collection strategy was executed from a realist perspective to capture the decisions made by the course designers; however, it did not incorporate a relativist perspective with regard to stakeholders, through further qualitative investigation. This decision was made to avoid interference in course delivery. To control biased selectivity and reporting bias, the data were sourced through multiple sources, including finance logs (and notes), data submitted to the employer, the funder, and timesheets. A traceability log was maintained linking the study questions to the relevant data sources and the study findings.

Analyze

Data analysis centered on the 3 cost categories and followed the 7-step process for cost definition.

Category A: Concept and Measurement of Costs

The preproduction budget was analyzed for the following ingredient categories: (1) personnel, (2) estate charges, (3) equipment and materials, (4) indirect costs and (5) stakeholder costs. The initial budgets did not reflect time for stakeholder costs (effort from third-party lecturers); therefore, this was captured as the additional time that was monitored in the study (and added for budget variance calculation), as there was no value for this in the data submitted to the funder.

Category B: Placing Values on Ingredients

With the full cost of production defined, values were associated with each ingredient subcategory to reflect the chargeable cost (including direct and indirect costs).

Category C: Calculating Costs

As each course was implemented in 1 year, and the courses were Web-based, there were no multiyear costs to calculate; the one-time cost of the project and the variance of the projected budget to the actual budget were the only variables under consideration. To accomplish this, the variance calculation of the budget to the incurred costs was undertaken at the completion of the project. The variance calculation compares actual costs to adjusted standard conditions based on occurrence [28].

The variance calculation formula is as follows: Variance = Actual spending − Budgeted spending.

Analyzing Costs of Observed Budget Variance Calculations

To determine the reasons for favorable or negative budget variance, the course designers were interviewed to determine the factors contributing to budget variance. This qualitative work was planned via the consolidated criteria for reporting qualitative research [29] to ensure that the appropriate trained staff conducted interviews, study design included the purposeful sampling of the course designers, sessions could be validated in the interviews, and the resultant analysis and findings would be repeatable [29]. The sessions were conducted as semistructured interviews transcribed and coded using thematic analysis [30] using total quality management (TQM) as coding criteria. TQM [31] is a quality appraisal method used to analyze factors affecting operational efficiency [32]. TQM provides a means to categorize issues relating to people, process, or technology through applying a systems approach to management (see Figure 1). For each area of cost variance, the course designers were asked to review budget reports to identify stages in the project lifecycle for variances in forecast and to describe the contributing factors. After the interview, these were coded

XSL•FO
RenderX

independently by 2 researchers to create a novel means of interpreting the cost calculation variance. For example, if a cost variance was attributed to stakeholder costs, the researchers would examine reported quarterly budgets (or at the project time interval) and determine where the variance began occurring. If the variance commenced during the build stage of the project, the project plan was analyzed, and questions surrounding the activities of the project were asked of the course designers to determine the root cause.

The key themes for the TQM analysis are presented in each case indicating the summary perspective of areas for improvement or efficiency in e-Learning budget creation.

Share

The findings of the variance calculation and the deductive-inductive interpretation of reasons leading to variation were presented in a case report to the course design and production team. Feedback was gathered on analysis and results; the key findings for each report were prepared for publication for a peer-review journal.

Cross-Case Synthesis

To derive results from the composite analysis of the cases, this study makes use of cross-case study synthesis [13] as illustrated in Figure 2. The standard variables in the cases are centered on ingredients and their incurred cost variance from budget.

Figure 1. Isolating variance during project stage to total quality management criteria.

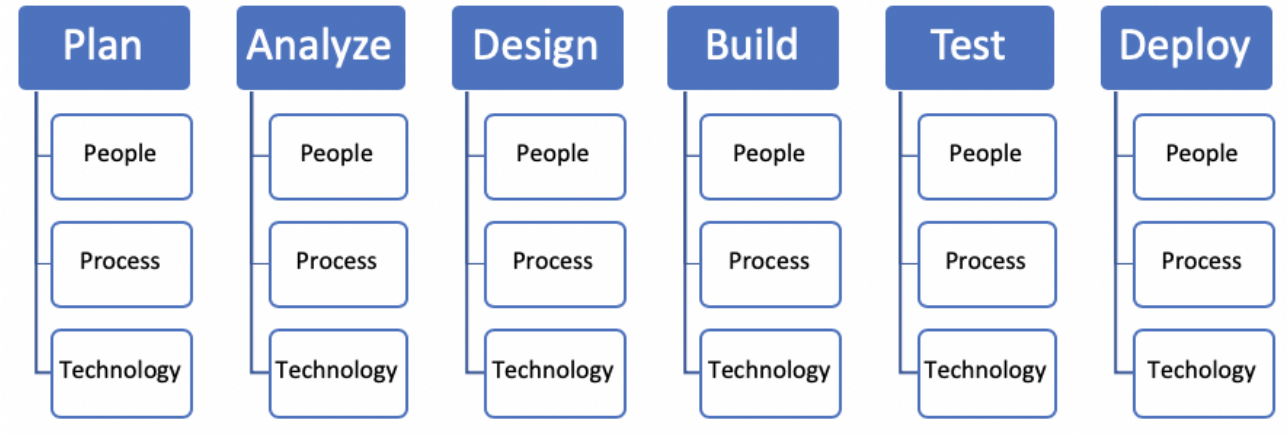
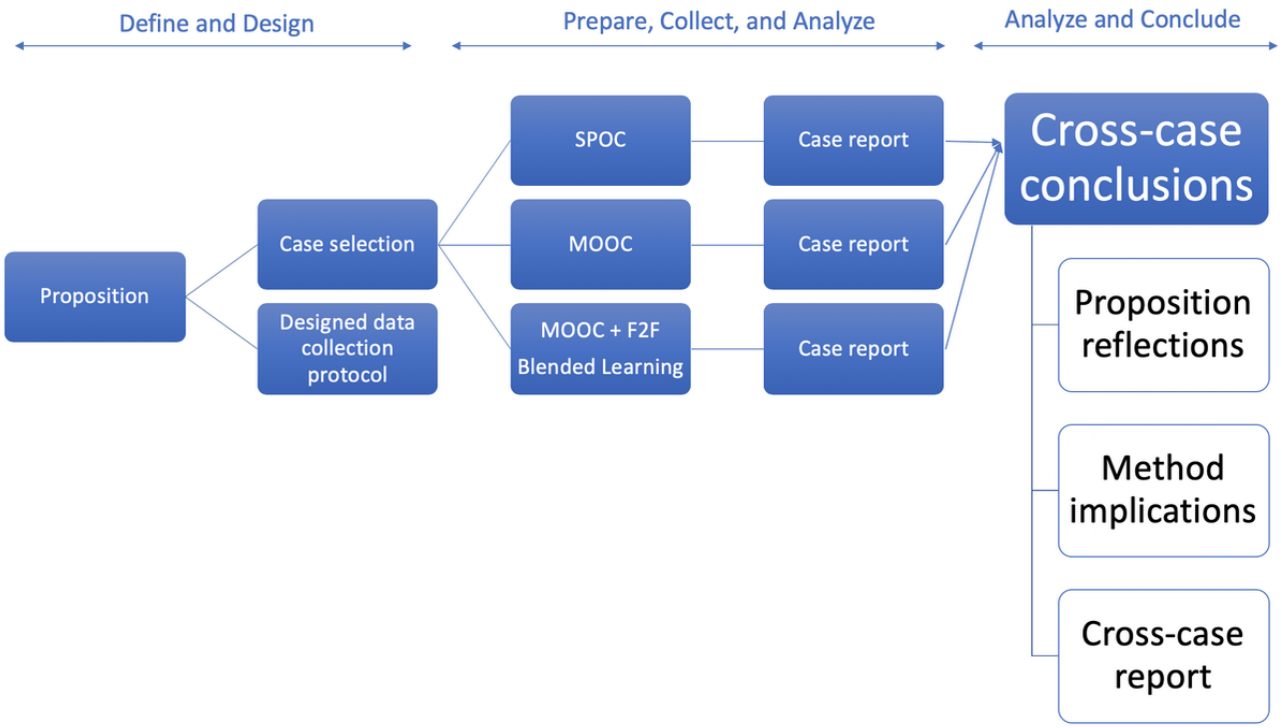


Figure 2. Cross-case synthesis. MOOC: massive open online course; SPOC: small private online course.



XSL•FO
RenderX

Results

Course Production Costs

Category A: Concept and Measurement of Costs

Costs for each case were summarized into components and separated into ingredient cost categories (Table 4).

Category B: Placing Values on Ingredients

Upon completion of the analysis of the ingredients of the course production, initial budgets were created and submitted to the funder.

Category C: Analyzing costs

Budget Variance Calculation

Case 1

The project implementation costs, in this case, had a negative variance of 41% (Multimedia Appendix 4). The most significant negative variance (135%; Multimedia Appendix 4) was in equipment and materials, primarily from the costs of app development in the creation of a Web-based course. As the production team had not created a Web-based course before, there was a significant underestimation of the amount of time required to build and configure the system (which was developed using the Open edX learning management system platform) and complete course editing. In addition, specialist recording equipment had to be procured that was not understood at the time of budget completion. The next most substantial negative variance (76%; Multimedia Appendix 4) was the amount of time required from third-party stakeholders in the production of learning materials. The amount of time allocated for recording the lecturers was underestimated; there had to be several re-runs of the recordings to address content changes. The lowest negative cost variance (31%; Multimedia Appendix 4) was in the personnel costs to deliver the course. Although the variance was the smallest of the 3 categories, it was significant because the course production team did not receive any additional compensation for their additional work; this extra work was captured in the project timesheets but not submitted to the funder for reimbursement.

Table 4. Ingredient categories.

Ingredient categories	Cost components
Personnel	University staff
Estate charges	Information technology services charges
Equipment and materials	Course production equipment and application development costs for the creation of software to support the massive open online course
Indirect costs	University overheads
Stakeholder costs	Staff for third-party subject matter consultancy

Case 2

The actual costs varied from the budgeted cost in personnel, equipment and materials, and stakeholder costs, and the total cost of production has a negative variance of 113% (Multimedia Appendix 4) from the budgeted amount. The most significant variance was in stakeholder costs, where the total time for external lecturers and subject matter experts to deliver work was significantly underbudgeted, with a negative variance of 190% (Multimedia Appendix 4). The reason for this underestimate was that videos had to be reshot twice and the amount of time allocated to retrieve stakeholders and complete associated course updates dramatically affected the budget. The second largest variance was in personnel; the cost variance was directly related to the additional production time required for the video reshoots, in addition to the iteration of the development of the platform. The course implementation online learning provider also switched from edX to FutureLearn learning management system during the project, requiring rework of previously completed tasks. As the team was not experienced on the FutureLearn platform, this further accounted for additional effort and the unfavorable budget variance; a team with experience and training on design for the course material would most likely have attained different results. Finally, equipment and materials were also underestimated with a negative variance of 133% (Multimedia Appendix 4), having to do with additional software required for video editing and additional workstations gathered to deal with additional editing required in the course development.

Case 3

In contrast to the previous case studies, this case demonstrated a positive variance of 16% (Multimedia Appendix 4) from the initial budget. Stakeholder costs for subject matter expert lecturers were slightly overestimated but close to budget. It is important to note that the third-party stakeholder team had significant previous experience working together for producing related coursework, and this could have led to the precision in effort estimation. Equipment and materials had a significant positive variance of 37% (Multimedia Appendix 4); the reason for this is that not all the equipment planned for the course development was necessary because there was efficiency derived in the course production and streamlining of data science modules that were thought to have required custom app development. Personnel had a negative variance of 13%; this was related to additional effort required in video editing. In addition, the course was completed ahead of schedule and in less time than was anticipated.

The construction of the cost ingredients and subsequent cost analysis underwent 3 validation tests (Table 5).

Issues affecting budget variance were classified using TQM to categorize factors influencing the budget (Table 6). Although each course was implemented with a varying form of

XSL•FO
RenderX

e-Learning, the issues affecting each case were similar and cross-applicable. The critical consideration in budgeting is less an aspect of the type of e-Learning, but more the planning associated with the project management of the creation of the course.

Table 5. Cross-case results validation tests.

Case	Construct validity	External validity	Reliability
1	To achieve data triangulation, the case study had multiple sources of cost data. (1) The project budget that was submitted to the project funder, (2) the actual costs submitted to the funder at the completion of the project, and (3) the timesheet log of hours captured by the course implementers. The final case report was reviewed, and feedback gathered from the course designers (BS, MT); any inconsistencies or inaccuracies were corrected.	By using Levin's ingredients method for cost identification, the case followed an established costing procedure that is used as the basis for analytic frameworks for economic evaluation in education. This process based on a common analytic framework allows for the generalization of the study findings to similar use cases.	A study protocol was created at the commencement of the case; the protocol details the structure of the study and details how data were collected to ensure the reliability of the results.
2	Multiple sources of cost data and reporting data were used to validate that data sources were an accurate record of what occurred. (1) The project budget created at the project commencement, (2) the actual cost report submitted at the completion of the project, (3) the timesheet log of hours captured by each team resource, (4) a third-party work-log for course production and monitor of billable hours recorded charged to the program, (5) external audit reports on the course construction, and (6) review of notes from monthly reviews of budget spend. The final case report was reviewed, and feedback gathered from the course designers (BS, MT); feedback was provided and reviewed by the research team to ensure implementation accuracy.	The repetition of a model used in prior research [22], application of Levin's ingredients method for education intervention analysis, and use of standard costing and variance calculation activity-based costing methods demonstrated a common analytic framework that is transportable to other studies.	To achieve this test, a study protocol was used and formed the governing basis for the study.
3	The data sources for each ingredient category were sourced from (1) the initial project budget, (2) reported submitted costs, (3) a time log of hours worked, and (4) a third-party work-log of the activities of subcontracted courses. The final case report was reviewed to ensure accuracy.	The same process that was used in the 2 previous cases was replicated [24], and application of Levin's ingredients method for education intervention analysis demonstrated a common analytic framework transportable to other electronic learning studies.	A minor variation of the previous study protocols executed was used and stored as the governance framework for the study.

Table 6. Total quality management category of issues affecting budget adherence to the model.

Cases	Issue	People	Process	Technology
Case 1	The inadequacy of project budgets at the commencement of Web-based learning for new teams	—[a]	X[b]	X
	Underreporting of personnel costs	X	X	—
Case 2	Resource task estimation and management	—	X	—
	Contingency planning	—	X	—
	Third-party resource management	X	X	—
	Need for an update of course materials	—	X	X
Case 3	Cost efficiencies in the delivery of a course piloted in previous years	—	X	—
	Experience and relationship of the course learning team	X	—	—
	Agile project management methods and iterative budget management	—	X	—

[a]Not applicable.

[b]Applicable.

Project Management

Each case implemented project management methods for the organization of crucial deliverables and tasks in their design and integrated learning design methodology in different ways. Case 1 employed project-related task-centered actions constructed to match each learning outcome. Case 2 integrated the analysis, design, development, implementation, and evaluation (ADDIE) model, and course planning was structured along each of these design stages, whereas case 3 implemented an agile project management model (with iterations) while using the ADDIE model in course construction.

Participant Information

Case 1

A total of 124 learners enrolled in the SPOC from September 2016 to December 2016 (Table 7). Of these, 84% completed

XSL•FO
RenderX

the course and received a postcourse certificate. The course uptake and completion, however, did not influence the production costs postcourse implementation as the course was designed as a self-managed SPOC not requiring further administration after deployment.

Table 7. Electronic learning implementation participation summary.

Case (year)	Learners, n	Completion, %
1: Educating administrative staff to engage with young patients (2016)	124	84
2: The impact of climate change on public health (2017) [19]	968	17
3: Data science in health care using real world evidence (2018)	5036	12

Case 2

A total of 968 learners participated in the MOOC from November 2017 to December 2017 (Table 7). Of these, 17% completed the course. The course completion ratio was in line with completion rates for MOOCs [33], where although there is a high uptake of initial learners, completion of course activity ranges from 8% to 20%.

Case 3

A total of 5036 learners participated in the MOOC from September 2018 to December 2018 (Table 7). Of these, 12% completed the course. The course completion ratio was also in line with completion rates for MOOCs [33]. A blended residential course was held in November 2018, with the participation of 14 learners (these learners were inclusive in the MOOC set). In this residential course, the participants completed the MOOC as prelearning and then undertook case studies, putting course learning into practice.

Discussion

Principal Findings

This study aimed to establish an approach for identifying the costs in the design, development, and deployment of applied health professions e-Learning courses. The standard components for the construction of an e-Learning course were determined by the methods used in this study, which combined existing approaches for cost budgeting with qualitative methods for the interpretation of results. Although Levin's ingredients method provides a mechanism for categorizing costs design and implementation costs for budgeting, TQM provides a qualitative framework to examine the effect of the design and production decisions on the budget. The key issues affecting the ability of the budget to deliver in line with expectations at the close of the project were related to process issues. Familiarization with technology was also a key issue in cases 1 and 2, where familiarity with production methods and learning technology had an effect on anticipated effort.

The key recommendations made from examination of these cases center on 3 areas of process-related enhancement, 1 having to do with project management and the remaining 2 having to do with budget management, both related to the course production and instructional design:

Project Management: Linkage of Instructional Design Method to Stages in the Project Lifecycle With Time Tracking

Project management enables the planning and prioritizing of activities; management of risk, issues, and actions; and ensuring quality. In these observed cases, the use of robust project management methods and the development of iterative methods to validate learning materials tended to create favorable results. In addition, linking an instructional design approach to project stages and tracking tasks by time to each component creates awareness and links the associated financial effect of delivery to course building.

Budget Planning: Use of Confidence Factors in Budget Time Estimating

A vital issue in all cases was overestimating the amount of effort required to build tasks. To better manage time tracking, we have suggested tracking task by time linked to learning design, but as an additional measure, building confidence factors into budgets allows a degree of error and contingency when building initial budgets. A confidence factor is a percentage of variance added to an initial cost forecast that can be added as a contingency; applying confidence factors based on requirements, the familiarity of approach, and other factors can lead to higher estimation precision.

Budget Planning: Modeling Budget Forecasting on Similar Implementations

Case 3 was the most successful in delivery because the course team had worked together delivering similar content, was able to gain efficiency in having preexisting relationships, and had an evidence base to build their cost models from. When planning e-Learning implementations, the starting point should similarly be previous projects or using data from the literature on factors influencing costs, so budgets are not determined from scratch. Part of the observed budget variance issues in cases 1 and 2 had to do with estimates for costs not built on prior evidence; this can be controlled by using an experience-driven starting point.

Strengths and Limitations

This study analyzes 3 distinct cases of e-Learning covering 6128 applied health learners in 3 years and provided a comprehensive summary of the issues affecting the production and development of a course. This information could be useful for course designers in the planning of their e-Learning implementations and for drawing on lessons learned to plan budgets that ensure projects meet their objectives.

XSL•FO
RenderX

We noted 4 limitations with this study. Case study research can only provide a snapshot of activities as observed in each case, and there is a possibility that these cases may have limited applicability to other contexts. This has been mitigated using construct validity, external validity, and reliability tests in each case, but it is important to note that case study research has an inherent limitation in the observation of events under consideration due to the design; experimental methods deliver more rigorous results to test results. In addition, the selection of the case studies was opportunistic, as they were e-Learning projects accessible within the first author's research unit. The second limitation is that further qualitative investigation of attitudes, views, and perceptions of stakeholders was not undertaken. This would have added an additional dataset to analyze factors affecting budgeting, meaning that the researchers drew conclusions from data that may have been viewed differently with further direct inquiry from stakeholders. It is important to note however, that stakeholders did review final case reports for accuracy and consistency with events. The third limitation is that the study did not undertake critical examination of the decisions made by the course designers in authoring tools, license costs, expertise, and other factors affecting the direct costs; examination of these costs including triangulation among the 3 sources would lead to further evidence affecting results. Finally, the study made use of a mixed-methods approach to analyze horizontal budget analysis but did not undertake an analysis for offsetting or magnifying variances, return on investment, forecasting, sensitivity analysis, or other financial planning and analysis methods. An economic study focused on outcomes and cost could provide further data that would potentially influence implementation considerations.

Further Research

The outputs of this study, in addition to the process of execution and reflection on both strengths and limitations, suggest 3 possible areas for future research:

Standards for Costing Economic Evaluations of e-Learning Implementations

Limited economic evaluations are conducted on e-Learning, most likely because educators focus on content delivery and educational effect rather than creating cost evidence. This study has created an extension of existing costing methods and demonstrated how it can be applied to e-Learning, allowing future researchers to reuse this approach to create consistent costing data, which could be subsequently benchmarked. With a growing evidence base of e-Learning cost data, this could also promote further research into various forms of economic evaluation, to create possible business cases for future investment in e-Learning, should value be demonstrated.

Integration of Project Management, Instructional Design Methods, and Costing

This study observed benefits in the combination of project management methods and instructional design methods; further research investigating ways of adopting existing instructional design methods with project management methodologies and linking these methods with cost management approaches could help address the high investment cost required in e-Learning.

Cost and Value Perceptions of Students and Educators

Using improved cost data from the approaches in this research, further research could attempt to identify perceptions of cost and value by comparing the perspectives of students and educators.

Conclusions

e-Learning research consistently refers to the promise and opportunity of its cost-effectiveness in contrast to face-to-face instruction; however, the underlying data supporting the costs necessary for their delivery are not well understood [8]. To implement further economic evaluation to understand proprieties demonstrating the value of e-Learning in contrast to other learning types, it is first necessary to develop a standard means of calculating costs in the delivery of these types of projects. Through consistent management of factors affecting costs in course production, further research could be undertaken using standard economic evaluation methods to evaluate the advantages of using e-Learning. This study enables an understanding of the issues affecting cost planning for the design, development, and deployment of e-Learning courses and also provides recommendations on controlling cost variance within e-Learning projects. This study contributes a systematic approach to costing in e-Learning that course designers and researchers could use to design and calculate costs in the production and deployment.

Acknowledgments

This project was supported by the European Institute of Innovation and Technology—EIT Health Knowledge and Innovation Community and the Higher Education Funding Council for England—Catalyst Fund. Boris Serafimov and Mel Toumazos provided detailed data contributing to the study in the design, development, and deployment of the e-Learning courses examined in this study. Yusuf Ermak and Hassan Chaudhury provided valuable contributions to the Data Science Courses as lecturers.

Authors' Contributions

EM conceived the study topic, wrote the first draft, responded to peer-review feedback, and is the principal investigator on the research project. DB, KF, and AA reviewed the completed draft manuscripts and provided feedback on iterations. JC supervised the investigation. EM is the guarantor.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1

Case study protocol – educating administrative staff to engage with young patients.

[DOCX File, 18KB - jmir_v21i6e13574_app1.docx]

Multimedia Appendix 2

Case study protocol – the impact of climate change on public health.

[DOCX File, 19KB - jmir_v21i6e13574_app2.docx]

Multimedia Appendix 3

Case study protocol – data science in healthcare using real world evidence.

[DOCX File, 18KB - jmir_v21i6e13574_app3.docx]

Multimedia Appendix 4

Ingredient costs variance calculation.

[DOCX File, 14KB - jmir_v21i6e13574_app4.docx]

References

1. Chen L, Evans D, Evans T, Sadana R, Stilwell B, Travis P, et al. Working together for health. In: Evans T, editor. The World Health Report 2006. Switzerland: Thomson Prentice; 2006:1-237.
2. Dakin H, Gray A. decision making for healthcare resource allocation: joint v separate decisions on interacting interventions. Med Decis Making 2018 Dec;38(4):476-486 [FREE Full text] [doi: 10.1177/0272989X18758018] [Medline: 29683792]
3. Al-Shorbaji N, Atun R, Car J, Majeed A, Wheeler E. In: Al-Shorbaji N, editor. eLearning for Undergraduate Health Professional Education - A Systematic Review Informing a Radical Transformation of Health Workforce Development. Switzerland: World Health Organisation; 2015:1-156.
4. Sangrà A, Vlachopoulos D, Cabrera N. Building an inclusive definition of e-learning: an approach to the conceptual framework. Int Rev Res Open Distr Learn 2012;13:146-159 [FREE Full text]
5. Alturkistani A, Majeed A, Car J, Brindley D, Wells G, Meinert E. data collection approaches to enable evaluation of a massive open online course about data science for continuing education in health care: case study. JMIR Med Educ 2019 Apr 2;5(1):e10982 [FREE Full text] [doi: 10.2196/10982] [Medline: 30938683]
6. Alturkistani A, Lam C, Foley K, Stenfors T, Blum E, Meinert E. MOOC evaluation methods: a systematic review. JMIR Preprints 2019:1-51 (forthcoming) [FREE Full text] [doi: 10.2196/preprints.13851]
7. Plint S. Pulse Today. NHS: NHS; 2014. Delivering the Mandate on GP Expansion (GP Taskforce Final Report) URL: http://www.pulsetoday.co.uk/download?ac=9243 [accessed 2019-05-19] [WebCite Cache ID 78TzIuDMV]
8. Meinert E, Reeves S, Eerens J, Banks C, Maloney S, Rivers G, et al. Exploring the cost of eLearning within the field of health professions education: key findings from a systematic scoping review. JMIR Preprints 2019:1-62 (forthcoming) [FREE Full text]
9. Hollands F, Tirthali D. Resource requirements and costs of developing and delivering MOOCs. Int Rev Res Open Distr Learn 2014 Oct 3;15(5):114-133. [doi: 10.19173/irrodl.v15i5.1901]
10. Alturkistani A, Car J, Majeed A, Brindley D, Wells G, Meinert E. Determining the Effectiveness of a Massive Open Online Course in Data Science for Health. : International Association for Development of the Information Society; 2018 Presented at: International Association for Development of the Information Society (IADIS) International Conference on e-Learning; July 17-19, 2018; Madrid, Spain URL: https://eric.ed.gov/?q=Determining+the+Effectiveness+of+a+Massive+Open+Online+Course+in+Data+Science+for+Health.&id=ED590297
11. Maloney S, Haas R, Keating J, Molloy E, Jolly B, Sims J, et al. Breakeven, cost benefit, cost effectiveness, and willingness to pay for web-based versus face-to-face education delivery for health professionals. J Med Internet Res 2012 Apr 2;14(2):e47 [FREE Full text] [doi: 10.2196/jmir.2040] [Medline: 22469659]
12. Hollin IL, Robinson KA. A scoping review of healthcare costs for patients with cystic fibrosis. Appl Health Econ Health Policy 2016 Apr;14(2):151-159. [doi: 10.1007/s40258-015-0211-4] [Medline: 26649564]
13. Yin R. Case Study Research and Applications: Design and Methods. Newbury Park, California: Sage Publications; 2017.
14. Creswell J, Hanson W, Clark Plano V, Morales A. Qualitative research designs: selection and implementation. Counsel Psychol 2016 Jun 30;35(2):236-264. [doi: 10.1177/0011000006287390]
15. Levin H, McEwan P, Belfield C, Bowden A, Shand R. Economic Evaluation in Education: Cost-Effectiveness and Benefit-Cost Analysis. Newbury Park, California: Sage Publications; 2017.

XSL•FO
RenderX

16. Kaplan R. Flexible budgeting in an activity-based costing framework. Account Horiz 1994;8(2):104-109 https://www.hbs.edu/faculty/Pages/item.aspx?num=2112 [FREE Full text]
17. Mak Y, Roush M. Managing activity costs with flexible budgets and variance analysis. Account Horiz 1996;10(3):141-146 [FREE Full text]
18. Mak Y, Roush M. Flexible budgeting and variance analysis in an activity-based costing environment. Account Horiz 1994;8(2):93-103 [FREE Full text]
19. Meinert E, Alturkistani A, Car J, Carter A, Wells G, Brindley D. Real-world evidence for postgraduate students and professionals in healthcare: protocol for the design of a blended massive open online course. BMJ Open 2018 Oct 4;8(9):e025196 [FREE Full text] [doi: 10.1136/bmjopen-2018-025196] [Medline: 30287615]
20. Reeves S, Pelone F, Harrison R, Goldman J, Zwarenstein M. Interprofessional collaboration to improve professional practice and healthcare outcomes. Cochrane Database Syst Rev 2017 Dec 22;6:CD000072 [FREE Full text] [doi: 10.1002/14651858.CD000072.pub3] [Medline: 28639262]
21. Belfield C, Bowden A, Levin H. Cost estimation in education: the ingredients method. In: Teaching Benefit-Cost Analysis. United Kingdom: Edward Elgar Publishing; 2018.
22. Meinert E, Eerens J, Banks C, Car J. A case study examining the cost measurements in production and delivery of a Small Private Online Course (SPOC) for teaching administrative staff to engage with young patients. 2019 Presented at: 11th annual International Conference on Education and New Learning Technologies; July 1-3, 2019; Palma de Mallorca, Spain.
23. Meinert E, Alturkistani A, Murray K, Sabatier P, Car J. A case study examining the cost measurements in production and delivery of Massive Open Online Course (MOOC) for teaching the relationship between health and climate change. : International Association for Development of the Information Society; 2019 Presented at: International Conference on e-Learning 2019; 2019; Porto, Portugal.
24. Meinert E, Alturkistani A, Foley K, Stenfors T, Chaudhury H, Ermak Y, et al. A case study examining the cost measurements in production and delivery of a blended massive open online course (MOOC) on the use of data science in healthcare using real world evidence. - 2019:- (forthcoming).
25. Kaplan R. Management accounting (1984-1994): development of new practice and theory. Manage Account Res 1994 Sep;5(3-4):247-260. [doi: 10.1006/mare.1994.1015]
26. Mak Y, Roush M. Managing activity costs with flexible budgeting and variance analysis. Account Horiz 1996;10(3):141-146 https://www.researchgate.net/publication/289652051_Managing_activity_costs_with_flexible_budgeting_and_variance_analysis.
27. Levin H. Waiting for Godot: cost - effectiveness analysis in education. New Direct Eval 2001;90:55-68 [FREE Full text] [doi: 10.1002/ev.12]
28. Drury C. Management and Cost Accounting. Boston, Massachusetts, United States of America: Cengage Learning; 2017.
29. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
30. Braun V, Clarke V. What can. Int J Qual Stud Health Well-being 2014;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
31. Lobo SR, Samaranayake P, Subramanian N. The impact of TQM and information communication technology (ICT) as an enabler in the quality management assessment framework (QMAF) on business outcomes. Int J Syst Sci Oper Logist 2017 Jul 4;6(1):69-85. [doi: 10.1080/23302674.2017.1344329]
32. Manzoor A. Quality assurance in open and distance learning. Optimiz Open Dist Learn High Educ Inst 2018:195-212 [FREE Full text]
33. Li Q, Wan F. A Case Study of the Characteristics of MOOCs Completers: Taking an Online Professional Training MOOC for Example. 2016 Presented at: 2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT); July 25-28, 2016; Austin, TX, USA p. 503. [doi: 10.1109/ICALT.2016.2]

Abbreviations

ADDIE: analysis, design, development, implementation, and evaluation
MOOC: massive open online course
SPOC: small private online course
TQM: total quality management
WHO: World Health Organization

XSL•FO
RenderX

Edited by G Eysenbach; submitted 01.02.19; peer-reviewed by D Baines, T Tanigawa, A Alharthi, J Lander; comments to author 28.02.19; revised version received 26.04.19; accepted 10.05.19; published 04.06.19.

Please cite as:
Meinert E, Alturkistani A, Foley KA, Brindley D, Car J
Examining Cost Measurements in Production and Delivery of Three Case Studies Using E-Learning for Applied Health Sciences: Cross-Case Synthesis
J Med Internet Res 2019;21(6):e13574
URL: https://www.jmir.org/2019/6/e13574/
doi:10.2196/13574
PMID:31165718

©Edward Meinert, Abrar Alturkistani, Kimberley A Foley, David Brindley, Josip Car. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Forecasting Implementation, Adoption, and Evaluation Challenges for an Electronic Game–Based Antimicrobial Stewardship Intervention: Co-Design Workshop With Multidisciplinary Stakeholders

Enrique Castro-Sánchez[1], MPH, PhD; Anuj Sood[1], BEng, MBA; Timothy Miles Rawson[1], MD, PhD; Jamie Firth[2]; Alison Helen Holmes[1], MD

[1]National Institute for Health Research Health Protection Research Unit in Healthcare-Associated Infection and Antimicrobial Resistance, Imperial College London, London, United Kingdom

[2]Jamie Firth Consultancy Ltd, London, United Kingdom

Corresponding Author:
Enrique Castro-Sánchez, MPH, PhD
National Institute for Health Research Health Protection Research Unit in Healthcare-Associated Infection and Antimicrobial Resistance
Imperial College London
London,
United Kingdom
Phone: 44 0203 3132732
Fax: 44 208 383 3394
Email: e.castro-sanchez@imperial.ac.uk

Abstract

Background: Serious games have been proposed to address the lack of engagement and sustainability traditionally affecting interventions aiming to improve optimal antibiotic use among hospital prescribers.

Objective: The goal of the research was to forecast gaps in implementation, adoption and evaluation of game-based interventions, and co-design solutions with antimicrobial clinicians and digital and behavioral researchers.

Methods: A co-development workshop with clinicians and academics in serious games, antimicrobials, and behavioral sciences was organized to open the International Summit on Serious Health Games in London, United Kingdom, in March 2018. The workshop was announced on social media and online platforms. Attendees were asked to work in small groups provided with a laptop/tablet and the latest version of the game On call: Antibiotics. A workshop leader guided open group discussions around implementation, adoption, and evaluation threats and potential solutions. Workshop summary notes were collated by an observer.

Results: There were 29 participants attending the workshop. Anticipated challenges to resolve reflected implementation threats such as an inadequate organizational arrangement to scale and sustain the use of the game, requiring sufficient technical and educational support and a streamlined feedback mechanism that made best use of data arriving from the game. Adoption threats included collective perceptions that a game would be a ludic rather than professional tool and demanding efforts to integrate all available educational solutions so none are seen as inferior. Evaluation threats included the need to combine game metrics with organizational indicators such as antibiotic use, which may be difficult to enable.

Conclusions: As with other technology-based interventions, deploying game-based solutions requires careful planning on how to engage and support clinicians in their use and how best to integrate the game and game outputs onto existing workflows. The ludic characteristics of the game may foster perceptions of unprofessionalism among gamers, which would need buffering from the organization.

(J Med Internet Res 2019;21(6):e13365) doi:10.2196/13365

KEYWORDS

serious games; antimicrobial stewardship; medical education

XSL•FO
RenderX

Introduction

The Threat of Drug-Resistant Infections

Although antimicrobial resistance is an evolutive phenomenon that cannot be stopped [1], reducing inadequate use of antibiotics would prolong their effectiveness and mitigate the clinical, human, and economic costs of drug-resistant infections [2]. As an example of such costs, in 2015 an estimated 672,000 infections with antibiotic-resistant bacteria were reported in the European Union [3], and yearly drug-resistant infection–attributable mortality worldwide has been forecasted to reach 10 million people by 2050 [4,5].

Antimicrobial stewardship (AMS) fosters the optimal use of antibiotics by health care professionals, patients, and citizens [6], combining organizational, structural, behavioral, and educational components. Several initiatives have reported on educational resources and interventions focused on undergraduate human health and veterinary students [7] as well as existing health care workers [8]. These interventions aimed to address existing gaps in undergraduate curricula [9], one of the factors responsible for suboptimal antibiotic prescribing practices reported worldwide [10].

Efforts to support existing and future antibiotic prescribers have been evidenced by the burgeoning number of educational resources already developed [11]. However, most of these resources have focused on improving the technical knowledge of prescribers about infections or antimicrobials [12]. While such education would be undoubtedly useful and of some benefit, it may have overlooked increasingly recognized behavioral influences on antimicrobial decision-making [13]. Surveys of trainee doctors and students evaluating educational interventions, for example, have highlighted how professionals still felt hesitant about their competence in antibiotic prescribing [14,15]. Other studies have stressed the inaction reported by some clinicians to modify antimicrobial prescriptions [16], demonstrating the communication, negotiation, and emotional skills required for antimicrobial decision-making [17]. It would be unlikely for such skills to be nurtured by passive approaches based on providing knowledge, instead requiring active and dynamic educational experiences that would allow the performance of trainees and clinicians to be examined and reviewed [18].

The combination of increasing computing power, ubiquity of portable devices, and near-complete internet coverage affords an ever-growing reliance on interventions based on games as clinical training aids and simulations [19]. Software development has also benefited from an increased understanding of the behavioral determinants of clinical decision-making and heuristics [20], with existing software now able to address some of the challenges presented by the frequent lack of engagement with traditional quality improvement interventions [21,22].

Recognizing the potential of game-based solutions to facilitate education and training of health care students and existing clinicians involved in antibiotic management, in 2015 we developed On call: Antibiotics (OcA), the first serious game worldwide aimed at improving antimicrobial prescribing behaviors among hospital prescribers [23] (Figure 1 and Multimedia Appendix 1).

Figure 1. On call: Antibiotics. Interface (left panel); hospital layout (right panel).

XSL•FO
RenderX

On Call: Antibiotics Development and Use

A close collaborative of artists, commercial game developers, health care workers, academics, and patient representatives came together to design the game. The platform resembles clinical practice and presents a number of virtual patients that require management using essential diagnostic skills and the broad range of optimal behaviors embedded in established national antibiotic guidance in the United Kingdom such as Start Smart Then Focus (SSTF) [24], which offers these principles of optimal antibiotic use:

- Do not start antibiotics in the absence of clinical evidence of bacterial infection.
- For antibiotics prescribed, document each of the following on the drug chart and in the clinical notes: clinical indication (including disease severity if appropriate), dose, route, and duration or review date.
- Obtain cultures first where possible.
- Prescribe single-dose antibiotics for surgical prophylaxis where antibiotics have been shown to be effective.
- Review the clinical diagnosis and continuing need for antibiotics by 48 to 72 hours and make a clear plan of action: the antimicrobial prescribing decision.
- The five antimicrobial prescribing decision options are stop, switch, change, continue, and refer to outpatient parenteral antibiotic therapy.
- It is essential that the review and subsequent decision be clearly documented in the clinical notes. The decision should also be documented clearly on the drug chart.

Players are presented with demographic information about each patient (Figure 1, left panel) together with a chief complaint that can be explored by requesting more information about signs and symptoms. Once players are ready to make a diagnostic decision, they select a diagnosis and a therapeutic option according to the SSTF principles. As soon as the therapeutic decision is made, players are given feedback about their performance from the perspective of the different professional groups involved in antibiotic decisions. In addition to such individual feedback, metrics such as time engaged with the game and each case and time needed to make each decision are collected by the game and submitted to a secure, remote server for analysis.

The game does not aim to teach clinicians about specific antibiotics and their appropriateness or effectiveness to treat a given infection on the premise that such antibiotic-microorganism combinations are likely to change in time and be context-specific and that technical solutions (ie, those based on the provision of information or support to address a priori knowledge deficits), albeit effective, have been short-lived and demonstrated modest impact [25]. Instead, we were more interested in fostering excellent antimicrobial behaviors, recognizing that such optimal nature may depend on a variety of interrelated professional, clinical, and organizational conditions [26] and patient expectations [27].

To reflect the range of demands exerted on the different stakeholders within the antimicrobial prescribing pathway, we developed some archetypes or personae that embodied the goals of the different stakeholders. For example, the ideal outcome for patients may merely be the resolution of their infection in the fastest possible way, disregarding the potential impact that such a forceful approach may have on the rates of drug-resistant infections in the wider population. Hospital prescribers (in this version of the game, doctors), on the other hand, may strive for diagnostic and therapeutic accuracy, underpinning the wider reputation of the clinical team or the institution. Regarding the hospital, managers may have concerns about the use of costly intravenous medications and the increased number of patients developing adverse events associated with the peripheral vascular devices required to administer such antibiotics.

By definition, these archetypes are unidimensional and, to a point, simplistic, but they help articulate the tensions faced by prescribers. However, the foundation of such multifaceted perspective was the idea that satisfying all relevant stakeholders along the antimicrobial decision-making process may not be possible, and that even adequate therapeutic and clinical decisions may lead to negative consequences and experiences for some stakeholders. Additionally, using the personae ensured that although practitioners in the United Kingdom were the intended audience of the current game, health care workers worldwide involved in management of antibiotics in other countries could feel represented.

From game release on October 1, 2015, to September 2018, there were approximately 4000 downloads with about 2100 unique game users worldwide (source: website analytics, Google Data Studio reports). Our dissemination approach has been cautious until now, showcasing the game as a tool with the potential to influence antimicrobial prescribing among clinicians but avoiding any claims of efficacy or effectiveness until the results of pending evaluations are obtained.

Our previous literature review of serious games for infection prevention and control and AMS had identified a handful of reports within this field [28]. In their reporting, the majority of software and apps emphasized development aspects followed by the theoretical underpinnings of the products. However, few if any papers offered a detailed account of implementation, adoption, or evaluation perspectives and challenges. Such absence of published experiences seems to widely affect the field of serious games in medical and clinical education. For example, Gorbanev [29] recently explored the quality of the evidence of game effectiveness for medical learning, noting a weakness of diffusion mechanisms and lack of repeated implementations of games in different settings, highlighting how in general games were implemented, tested, and reported only once and in only one setting, with no comparator. Additionally, the unintended consequences of incorporating serious games as educational and behavioral tools in clinical education, including game elements and techniques as a source of distraction, have not been fully explored [30]. Addressing this gap in the evidence remains crucial to ensure meaningful impact in the real world, particularly important considering that these serious games intend to modify behaviors.

For such reasons, and to help identify potential threats to the successful deployment, implementation, and adoption of the game, we organized a co-development workshop with clinicians

XSL•FO
RenderX

and academics in serious games and antimicrobial use to debate those gaps and co-design solutions.

Methods

Attendee Selection

The workshop was widely advertised on our institutional website, event-booking platforms, and social media. We reviewed all individuals registered to attend the event and rejected people who could not demonstrate clinical expertise in antibiotic use or experience developing serious games or researching the theoretical and behavioral frameworks included in serious games to justify their attendance. We added such evaluation to ensure that room capacity for the event was not surpassed. Ideal participants were involved in prescribing, reviewing, or administering antibiotics; developing serious and other types of games; or developing or researching behavioral methods or tools, regardless of whether they focused on digital or nondigital delivery approaches.

Purpose of the Workshop

To maximize participation and cost efficiency, the workshop took place prior to the International Summit on Serious Health Games [31] organized by the Health Protection Research Centre for Healthcare Associated Infection and Antimicrobial Resistance at Imperial College London on March 20, 2018, to showcase innovative serious health games that had demonstrated robust or imaginative methodologies during their development, implementation, or evaluation or focused on topics transferable to the purpose of our game.

The 2-hour workshop was arranged from a perspective of co-design and coproduction, although we recognized that participants would be presented with a fairly stable and defined version of the game, allowing no modification of either the user interface or functionality. The layout of the workshop venue was arranged so attendees could work in small groups around a laptop or tablet running the game following brief presentations from the workshop organizer (ECS). The organizer summarized responses from each small group, seeking their agreement about the fidelity of such summary and noting divergent opinions on flipcharts. Workshop participants could also download the latest game version on their personal devices from the institutional repository [32].

We urged attendees to hypothesize whether the use of the game could lead to unintended consequences or unexpected events. Reflecting the gaps that had been identified in the literature review, experts were asked to sequentially debate the challenges in implementation, adoption, and evaluation of OcA, as highlighted in the manuscript by Castro-Sanchez et al [28]. For example, among the adoption challenges, participants had to identify the strengths, weaknesses, opportunities, and threats for clinicians, end users, technologists, developers, and those responsible for funding or commissioning the game. Regarding implementations challenges, participants were specifically asked to reflect upon the scalability and sustainability of OcA as a learning and behavior modification tool for clinicians. Finally, the section on evaluation challenges requested that workshop participants identify qualitative, quantitative, mixed, and economic evaluation approaches from the point of view of clinician players and technologists or developers engaged in the evaluation, with ideas about the type of evidence funders and commissioners may wish to receive to approve implementation and adoption of the game across new organizations.

Participants were encouraged to consider perspectives ranging from the individual clinician to the organization and the health service. They were asked to consider potential challenges to translating the intervention to low-resource settings to allow it to run on virtually any computing platform and focus its design away from particular therapeutic regimens and toward behaviors that have been agreed upon in many settings already. Although the workshop preceded the Summit, attendance was voluntary.

Analysis

As the workshop had a coproduction perspective, the goal was to identify actionable solutions that could be iterated in real-world use of the game and refined rather than establishing hypotheses or exploring perceptions or ideas of the participants about the implementation of the game. Threats and solutions were summarized by the workshop leader from opinions expressed during the workshop and by each team of participants after each of the workshop sections. These summaries were corroborated or further clarified by the participants and collectively agreed. Notes from the observer also documented interactions between participants and emerging interesting ideas. This paper synthesizes threats and solutions generated during the workshop.

Results

Characteristics of Participants

Of the applicants, 29 were selected to attend the workshop, including consultant physicians from different clinical specialties; doctoral and postdoctoral researchers with projects focused on simulation, games, or virtual environments; antimicrobial resistance researchers and clinicians; experts in digital intervention implementation; games developers; behavioral researchers with interests in game-based interventions; and funding and digital project managers. Workshop members came from the United Kingdom, Spain, Germany, Portugal, France, and Ireland. Table 1 summarizes the solutions coproduced at the event.

XSL•FO
RenderX

Table 1. Solutions proposed by workshop attendees.

Stage	Threat	Solution
Implementation	Inadequate organizational set-up for scalability and sustainability (technical and educational support for players and others, data handling, workforce volatility)	Consult practice educators and tutors on future game versions to preserve professional and game pedagogies aligned
Adoption	Perceptions about games as ludic rather than serious tools; questioning professionalism of game users	Identification of optimal game users and timing. Embedding game in other training and education solutions
Evaluation	Unsatisfactory evaluation frameworks so added value of games to existing multimodal bundles cannot be robustly measured	Multiple mixed-methods evaluations (quantitative, qualitative, and in-game) that provide added value to existing tools; identification and linkage to in-practice and cost-relevant metrics

Implementation Challenges

The most pressing implementation challenges identified by participants referred to the organizational setup necessary to ensure scalability and sustainability of the game as a viable professional tool. For participants, successful implementation of OcA would require the release of the game underpinned by robust technical and educational support for players, with inclusion of other nonplaying clinicians and any other professionals in education in practice.

Ideally, these educationalists or practice educators and tutors would remain consulted about any future game iterations so required learning or professional outcomes are reflected in the game. Similarly, education and training leads would receive intelligence from in-game metrics and a variety of proposed evaluation results to maintain a coherent and aligned pedagogy.

The scalability of the game as a platform was not really considered to be a threat to the success of the intervention, as the total potential number of gamers at the institution would not be overwhelming to the resources in place to host the software. Instead, attendees felt that, potentially, the volume of data generated by the game itself combined with clinical information and experiential gamer feedback may end up being overwhelming for organizational follow-up and lack sufficient granularity to enable continued practice improvement. As an example, attendees wondered how it would be possible to maintain personalized feedback, training, and education for each of the players, potentially most medical prescribers in the hospital.

Linked to the previous concern but in terms of sustainability, the cyclical nature of the UK trainee medical workforce targeted, transferring from one health care organization to another every few months, may obfuscate the meaningful follow-up of players. What should be done, for example, when clinicians leave the organization yet remain engaged with the software to avoid their in-game generated data confuse decisions by the hospital management?

Adoption Challenges

Attendees agreed that adoption of a game-based intervention such as OcA as an adjuvant to support clinical practice may suffer from perceptions about serious games as valid behavioral tools. As suggested previously, potential end users may reject the tool simply on the basis of their ideas about a game as a ludic experience and not the proper training required to address the assumed lack of knowledge about antibiotic guidelines, considered to be the ultimate reason for inappropriate prescribing.

Participants added that to encourage game adoption it would be necessary to make explicit who should be using the game and the timing of its use. Such unambiguous guidance would also be welcomed as OcA, albeit reflecting clinical realities, was not a simulator (ie, a life-like representation), like other tools such as surgical training platforms, or readily available, like advanced life support stations.

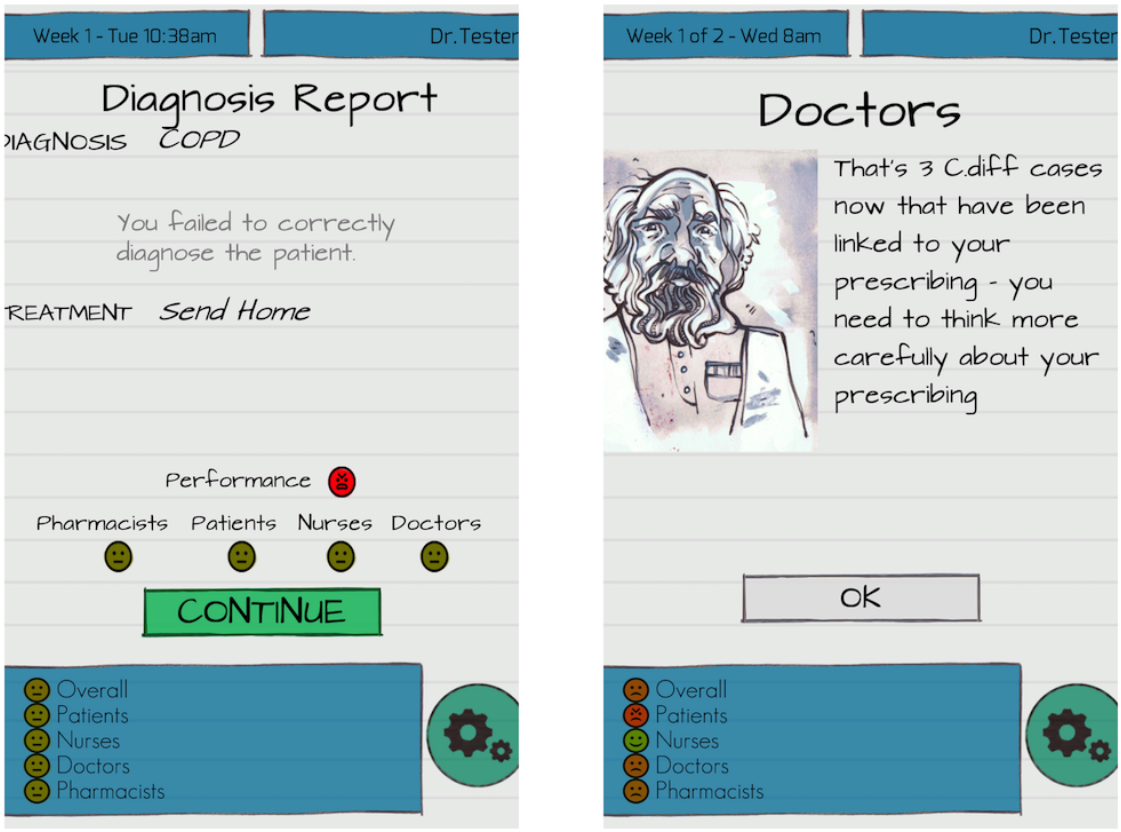
Regarding the characteristics of the ideal candidate players, clinicians unfamiliar with the demographics of gaming may assume that games would just be appealing or valid for, generally speaking, male and younger health care professionals. These perceptions, however, may not be aligned with existing evidence. Regarding the timing of game use, the workshop attendees suggested that potential end players may hesitate to be seen engaging with the game during working hours for fears of doubts about their professionalism. Such perceptions should not be dismissed and could be robust enough to derail the sustained adoption of OcA and similar tools.

To mitigate such threats, attendees advocated for an adequate communication campaign about the evidence underpinning the use of the game and the inclusion of the game within the pool of training and continuous professional development interventions offered by the organization. For instance, there should be concerted efforts to integrate this game-based platform among established performance review mechanisms such as portfolios, embedding game scores and feedback or using in-game performance as reflection to be jointly discussed between assessors and appraisees to evaluate skills and decision-making (Figure 2).

Interestingly, some attendees highlighted the dilemmas that may arise when proposing the use of games as interventions aimed at resolving attrition and lack of engagement. Although unlikely, it was hypothesized that extremely successful initiatives may nudge end users to play too much or at inappropriate times.

XSL•FO
RenderX

Figure 2. In-game feedback. Immediate performance (left panel); cumulative performance (right panel).



Evaluation Challenges

Attendees agreed on the clear need to ensure that game-based behavioral tools are robustly evaluated, ideally using different metrics and approaches, if they are to be postulated as effective and worthy of adoption by clinicians and funding by commissioners and decision-makers. Ideally, evaluations should be based on the most stringent yet feasible research design and focused on matching the game against current educational interventions or in addition to such interventions so the added value to existing multimodal bundles can be measured.

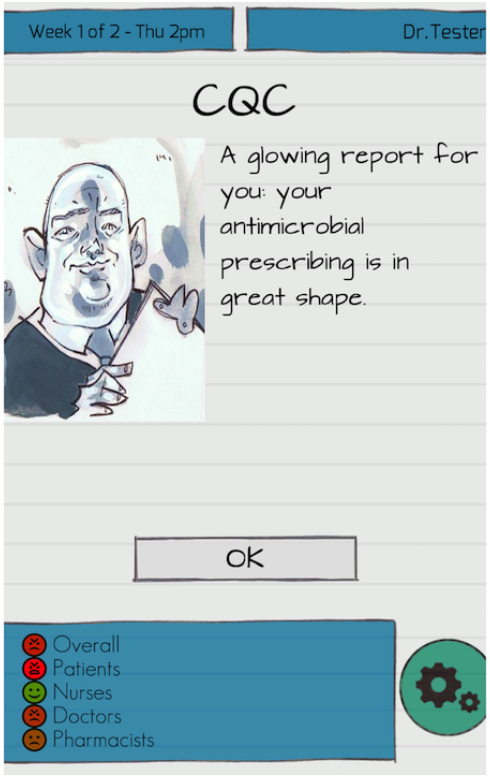
In addition to outcome-oriented evaluations, participants stressed the benefit of qualitative or mixed-methods designs that facilitated discussions with potential players about their embedding of serious games within professional or personal workflows, synergies with existing education or training approaches, and sustainability of perceived behavior changes.

For attendees, evaluation design depended on the niche identified for the game. As the game aspired to improve suboptimal prescribing behaviors, it would therefore be crucial to consider with care which outcome evaluation indicators to appoint. Further, as OcA introduced multiple perspectives regarding the impact of prescribers' decisions, agreeing on what optimal prescribing represented may be complex (Figure 3). Of interest, planned OcA evaluation activities include randomized controlled trials complemented by semistructured interviews triangulated with in-game metrics.

Two final notes of caution were mentioned during the workshop. First, there were concerns about how to best link behavioral and clinical metrics regarding outcome evaluation. As in-game performance had to be matched to clinical, real-world performance, there would be a need to establish how individual behaviors in the game would be linked to performance reflected on clinical records. Obviously, this challenge would only be relevant if the unit of analysis for the adopted evaluation design focused on individual clinicians; should other designs center on wards or teams, other metrics may be preferable or suitable. Second, there were concerns about how to determine any cost-related benefits in view of the complexity to arrive confidently at game-attributable improvements in clinical practice.

XSL•FO
RenderX

Figure 3. In-game behavioral nudge.



Discussion

Principal Findings

The potential of innovative game-based interventions in clinical practice can be threatened by perceptions about the validity of such behavior change approaches, as reported by the attendees to our workshop. To mitigate such threat, game advocates ought to include explicit guidance and information in their implementation strategies about intended end players and when they could use the game. Ideally, games should be positioned as equal tools among other interventions, with efforts to build synergies between tools and approaches and, ultimately, work toward a multichanneled ecosystem of educational experiences.

To achieve such parity with traditional resources, our attendees recommended that games should undergo robust evaluations from multiple perspectives with the aspiration to achieve noninferiority against already established resources. Evaluations should focus on clinical rather than statistical success. The emphasis on real-world improvement would facilitate the assessment of cost-associated measures in order to convince decision-makers of the benefits of deploying the game.

Limitations

The main limitations affecting the results refer to the very circumscribed software discussed at the workshop, a serious game focused on improving clinical behaviors among prescribers in the United Kingdom; the moderate number of attendees to the workshop; and the format of the event, where opinions where sought and had to be offered openly following a brief period of contact with the game, which may have contributed to socially desirable responses. Other academics or software coders interested in developing game-based solutions to influence clinical behaviors may still benefit from our findings, particularly the prominence given to adequate evaluation designs, including qualitative approaches.

Comparison With Prior Work

Beside OcA, few if any serious game-based approaches have been published on antimicrobial stewardship. However, our workshop centered on mitigating any expected or foreseen implementation, adoption, and evaluation challenges likely to arise when deploying a game-based intervention in clinical practice.

From that perspective, our results complement work about technology implementation [33] by suggesting that organizations interested in games would have to dispel concerns about sensitivities on their appropriateness as valid educational instruments, affecting the brand attitude [34] and modulating the intention to use factor predicated within models such as the technology acceptance model [35] or technology integration model [36].

Further, although perceived ease of use has been recognized as a contributing factor toward games adoption, our results suggest that it would be paramount to consider such ease not just from the game itself but also from the embedment of the game within the existing continuous professional development and educational workflows gamers may have already in place. As OcA can link to professional portfolios and output gaming hours and performance scores easily, the integration into the workflow may not be that difficult, which may not be the case for every game and would require careful planning [37].

Finally, our attendees firmly endorsed robust and multiple evaluation approaches, a concern well described previously [38]. However, they stressed the economic aspects necessary

XSL•FO
RenderX

to convince decision-makers about the use of the game, perhaps reflecting wider difficulties of introducing novel technologies onto clinical practice and education at times of funding restrictions.

Conclusions

Game-based interventions can aid efforts to improve antimicrobial use, but their successful deployment and sustained use cannot rely on assumptions about any inherent interest to clinicians. Even if accurate, such assumptions would need to resolve perceptions about the use of games as professional learning platforms and be supported by organizational planning and multipronged metrics of success.

Acknowledgments

The research was funded by the National Institute for Health Research (NIHR) Health Protection Research Unit (HPRU-2012-10047) in Healthcare-Associated Infection and Antimicrobial Resistance at Imperial College London in partnership with Public Health England. The workshop was funded by an Economic and Social Research Council Impact Accelerator Award (ES/M500562/1). ECS has received a Wellcome Institutional Strategic Support Fund Faculty postdoctoral fellowship and an Early Career Research Fellowship from the Antimicrobial Research Collaborative at Imperial College London and acknowledges the support of the Florence Nightingale Foundation and the NIHR Imperial Patient Safety Translational Research Centre. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, the Department of Health, or Public Health England. The funders had no role in the conduct of the study, the writing of the manuscript, its approval for publication, or submission to the journal.

Authors' Contributions

ECS led the clinical development of the game, organized the implementation workshop, and led the writing of the manuscript. AS provided technical advice and reviewed the manuscript. TMR provided technical advice and reviewed the manuscript. JF led the technical development of the game, obtained game download data, and reviewed the manuscript. AHH provided technical advice and reviewed the manuscript. We are grateful to participants for their insights.

Conflicts of Interest

None declared.

Multimedia Appendix 1

“On call: Antibiotics”, a first antimicrobial stewardship behaviour change serious game.

[MP4 File (MP4 Video), 60MB - jmir_v21i6e13365_app1.mp4]

References

1. Holmes AH, Moore LSP, Sundsfjord A, Steinbakk M, Regmi S, Karkey A, et al. Understanding the mechanisms and drivers of antimicrobial resistance. Lancet 2016 Jan;387(10014):176-187. [doi: 10.1016/S0140-6736(15)00473-0]
2. Bell BG, Schellevis F, Stobberingh E, Goossens H, Pringle M. A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance. BMC Infect Dis 2014 Jan 09;14:13 [FREE Full text] [doi: 10.1186/1471-2334-14-13] [Medline: 24405683]
3. Cassini A, Högberg LD, Plachouras D, Quattrocchi A, Hoxha A, Simonsen GS, Burden of AMR Collaborative Group. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. Lancet Infect Dis 2019 Jan;19(1):56-66 [FREE Full text] [doi: 10.1016/S1473-3099(18)30605-4] [Medline: 30409683]
4. O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. The review on antimicrobial resistance. 2016 May. URL: https://amr-review.org/sites/default/files/160518_Final%20paper_with%20cover.pdf [accessed 2019-05-11] [WebCite Cache ID 78ICHZiXG]
5. Gandra S, Barter DM, Laxminarayan R. Economic burden of antibiotic resistance: how much do we really know? Clin Microbiol Infect 2014 Oct;20(10):973-980 [FREE Full text] [doi: 10.1111/1469-0691.12798] [Medline: 25273968]
6. Dyar O, Huttner B, Schouten J, Pulcini C. What is antimicrobial stewardship? Clin Microbiol Infect 2017 Nov;23(11):793-798. [doi: 10.1016/j.cmi.2017.08.026]
7. Courtenay M, Lim R, Castro-Sanchez E, Deslandes R, Hodson K, Morris G, et al. Development of consensus-based national antimicrobial stewardship competencies for UK undergraduate healthcare professional education. J Hosp Infect 2018 Nov;100(3):245-256. [doi: 10.1016/j.jhin.2018.06.022]
8. Rawson TM, Moore LSP, Gilchrist MJ, Holmes AH. Antimicrobial stewardship: are we failing in cross-specialty clinical engagement? J Antimicrob Chemother 2015 Oct 23;71(2):554-559. [doi: 10.1093/jac/dkv337]

XSL•FO
RenderX

9. Castro-Sánchez E, Drumright LN, Gharbi M, Farrell S, Holmes AH. Mapping antimicrobial stewardship in undergraduate medical, dental, pharmacy, nursing and veterinary education in the United Kingdom. PLoS ONE 2016 Feb 29;11(2):e0150056. [doi: 10.1371/journal.pone.0150056]
10. Knox MC, Edye M. Educational antimicrobial stewardship intervention ineffective in changing surgical prophylactic antibiotic prescribing. Surg Infect 2016 Apr;17(2):224-228. [doi: 10.1089/sur.2015.194]
11. Rogers Van Katwyk S, Jones SL, Hoffman SJ. Mapping educational opportunities for healthcare workers on antimicrobial resistance and stewardship around the world. Hum Resour Health 2018 Dec 05;16(1):1-18 [FREE Full text] [doi: 10.1186/s12960-018-0270-3] [Medline: 29402327]
12. Dyar O, Beović B, Pulcini C, Tacconelli E, Hulscher M, Cookson B, ESCMID Generic Competencies Working Group. ESCMID generic competencies in antimicrobial prescribing and stewardship: towards a European consensus. Clin Microbiol Infect 2019 Jan;25(1):13-19 [FREE Full text] [doi: 10.1016/j.cmi.2018.09.022] [Medline: 30414817]
13. Charani E, Castro-Sanchez E, Sevdalis N, Kyratsis Y, Drumright L, Shah N, et al. Understanding the determinants of antimicrobial prescribing within hospitals: the role of prescribing etiquette. Clin Infect Dis 2013 Jul;57(2):188-196 [FREE Full text] [doi: 10.1093/cid/cit212] [Medline: 23572483]
14. Dyar OJ, Pulcini C, Howard P, Nathwani D, Nathwani D, Beovic B, et al. European medical students: a first multicentre study of knowledge, attitudes and perceptions of antibiotic prescribing and antibiotic resistance. J Antimicrob Chemother 2013 Nov 04;69(3):842-846. [doi: 10.1093/jac/dkt440]
15. Navarro-San Francisco C, Del Toro MD, Cobo J, De Gea-García JH, Vañó-Galván S, Moreno-Ramos F, et al. Knowledge and perceptions of junior and senior Spanish resident doctors about antibiotic use and resistance: results of a multicenter survey. Enferm Infecc Microbiol Clin 2013 Apr;31(4):199-204. [doi: 10.1016/j.eimc.2012.05.016] [Medline: 22819389]
16. Kirby E, Broom A, Gibson A, Broom J, Yarwood T, Post J. Medical authority, managerial power and political will: a Bourdieusian analysis of antibiotics in the hospital. Health (London) 2017 Jun 24;22(5):500-518. [doi: 10.1177/1363459317715775]
17. Lum EP, Page K, Whitty JA, Doust J, Graves N. Antibiotic prescribing in primary healthcare: dominant factors and trade-offs in decision-making. Infect Dis Health 2018 Jun;23(2):74-86. [doi: 10.1016/j.idh.2017.12.002]
18. McCullough AR, Rathbone J, Parekh S, Hoffmann TC, Del Mar CB. Not in my backyard: a systematic review of clinicians' knowledge and beliefs about antibiotic resistance. J Antimicrob Chemother 2015 Sep;70(9):2465-2473. [doi: 10.1093/jac/dkv164] [Medline: 26093375]
19. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. Br J Surg 2012 Oct;99(10):1322-1330. [doi: 10.1002/bjs.8819] [Medline: 22961509]
20. Jones M, Butler J, Graber CJ, Glassman P, Samore MH, Pollack LA, et al. Think twice: a cognitive perspective of an antibiotic timeout intervention to improve antibiotic use. J Biomed Inform 2017 Jul;71:S22-S31. [doi: 10.1016/j.jbi.2016.06.005]
21. Molnar A. Antimicrobial resistance awareness and games. Trends Microbiol 2019 Jan;27(1):1-3. [doi: 10.1016/j.tim.2018.09.007]
22. Wang R, DeMaria S, Goldberg A, Katz D. A systematic review of serious games in training health care professionals. Simul Healthc 2016 Feb;11(1):41-51. [doi: 10.1097/SIH.0000000000000118] [Medline: 26536340]
23. Castro-Sánchez E, Charani E, Moore L, Gharbi M, Holmes A. “On call: antibiotics” - development and evaluation of a serious antimicrobial prescribing game for hospital care. In: Schouten B, Fedtke S, Schijven M, Vosmeer M, Gekker A, editors. Games for Health. Wiesbaden: Springer-Verlag; 2014.
24. Ashiru-Oredope D, Sharland M, Charani E, McNulty C, Cooke J, ARHAI Antimicrobial Stewardship Group. Improving the quality of antibiotic prescribing in the NHS by developing a new Antimicrobial Stewardship Programme: Start Smart–Then Focus. J Antimicrob Chemother 2012 Jul;67 Suppl 1:i51-i63. [doi: 10.1093/jac/dks202] [Medline: 22855879]
25. Bleasdale SC, Barnden M, Barnes S. The impact of antibiotic stewardship program resources on infection prevention programs. Clin Infect Dis 2018 Nov 20:1. [doi: 10.1093/cid/ciy986] [Medline: 30462184]
26. Morley GL, Wacogne ID. UK recommendations for combating antimicrobial resistance: a review of "antimicrobial stewardship: systems and processes for effective antimicrobial medicine use". Arch Dis Child Educ Pract Ed 2017 Aug 09:46-49. [doi: 10.1136/archdischild-2016-311557] [Medline: 28794004]
27. Zanichelli V, Tebano G, Gyssens I, Vlahović-Palčevski V, Monnier A, Stanic Benic M, et al. Patient-related determinants of antibiotic use: a systematic review. Clin Microbiol Infect 2019 Jan;25(1):48-53. [doi: 10.1016/j.cmi.2018.04.031]
28. Castro-Sánchez E, Kyratsis Y, Iwami M, Rawson TM, Holmes AH. Serious electronic games as behavioural change interventions in healthcare-associated infections and infection prevention and control: a scoping review of the literature and future directions. Antimicrob Resist Infect Control 2016;5:34 [FREE Full text] [doi: 10.1186/s13756-016-0137-0] [Medline: 27777755]
29. Gorbanev I, Agudelo-Londoño S, González RA, Cortes A, Pomares A, Delgadillo V, et al. A systematic review of serious games in medical education: quality of evidence and pedagogical strategy. Med Educ Online 2018 Dec;23(1):1438718 [FREE Full text] [doi: 10.1080/10872981.2018.1438718] [Medline: 29457760]
30. Maheu-Cadotte M, Cossette S, Dubé V, Fontaine G, Mailhot T, Lavoie P, et al. Effectiveness of serious games and impact of design elements on engagement and educational outcomes in healthcare professionals and students: a systematic review

XSL•FO
RenderX

and meta-analysis protocol. BMJ Open 2018 Mar 16;8(3):e019871. [doi: 10.1136/bmjopen-2017-019871] [Medline: 29549206]
31. International Summit on Serious Health Games. URL: https://www.imperial.ac.uk/news/184527/international-summit-serious-health-games/ [accessed 2019-05-10] [WebCite Cache ID 75H5p8qHg]
32. Antibiotic prescribing game.: NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance URL: https://www.imperial.ac.uk/medicine/hpru-amr/applications-and-tools/antibiotic-prescribing-game/ [accessed 2019-01-10] [WebCite Cache ID 75JsuilfZ]
33. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
34. Baptista G, Oliveira T. Gamification and serious games: a literature meta-analysis and integrative model. Comput Hum Behav 2019 Mar;92:306-315. [doi: 10.1016/j.chb.2018.11.030]
35. Malaquias RF, Malaquias FF, Hwang Y. Understanding technology acceptance features in learning through a serious game. Comput Hum Behav 2018 Oct;87:395-402. [doi: 10.1016/j.chb.2018.06.008]
36. Shaw H, Ellis DA, Ziegler FV. The Technology Integration Model (TIM): predicting the continued use of technology. Comput Hum Behav 2018 Jun;83:204-214. [doi: 10.1016/j.chb.2018.02.001]
37. Morschheuser B, Hamari J, Koivisto J, Maedche A. Gamified crowdsourcing: conceptualization, literature review, and future agenda. Int J Hum-Comput Stud 2017 Oct;106:26-43. [doi: 10.1016/j.ijhcs.2017.04.005]
38. Mayer I, Bekebrede G, Harteveld C, Warmelink H, Zhou Q, van Ruijven T, et al. The research and evaluation of serious games: toward a comprehensive methodology. Br J Educ Technol 2013 May 28;45(3):502-527. [doi: 10.1111/bjet.12067]

Abbreviations

AMS: antimicrobial stewardship
NIHR: National Institute for Health Research
OcA: On call: Antibiotics
SSTF: Start Smart Then Focus

Edited by G Eysenbach; submitted 10.01.19; peer-reviewed by C Moore, K Glover; comments to author 21.02.19; revised version received 26.03.19; accepted 09.04.19; published 04.06.19.

Please cite as:
Castro-Sánchez E, Sood A, Rawson TM, Firth J, Holmes AH
Forecasting Implementation, Adoption, and Evaluation Challenges for an Electronic Game–Based Antimicrobial Stewardship Intervention: Co-Design Workshop With Multidisciplinary Stakeholders
J Med Internet Res 2019;21(6):e13365
URL: https://www.jmir.org/2019/6/e13365/
doi:10.2196/13365
PMID:31165712

©Enrique Castro-Sánchez, Anuj Sood, Timothy Miles Rawson, Jamie Firth, Alison Helen Holmes. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predictors of Patients' Intention to Interact With Doctors in Web-Based Health Communities in China: Cross-Sectional Study

Tailai Wu[1], PhD; Zhaohua Deng[1], PhD; Zhuo Chen[2,3], PhD; Donglan Zhang[2], PhD; Ruoxi Wang[1], PhD; Xiang Wu[1], PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]Department of Health Policy and Management, College of Public Health, University of Georgia, Athens, GA, United States

[3]School of Economics, Faculty of Humanities and Social Sciences, University of Nottingham Ningbo China, Ningbo, China

Corresponding Author:
Xiang Wu, PhD
School of Medicine and Health Management
Tongji Medical College
Huazhong University of Science and Technology
13 Hangkong Road
Wuhan,
China
Phone: 86 13071253919
Email: wuhsiang@hust.edu.cn

Abstract

Background: Web-based health communities provide opportunities for doctors and patients to interact with each other and change the traditional communication mode between doctors and patients. However, little is known about the predictors of patients' intention to interact with doctors in Web-based health communities in China.

Objective: The purpose of this study was to investigate what are the predictors of patients' intention to interact with doctors in Web-based health communities in China.

Methods: On the basis of two-factor theory and service convenience theory, we propose that the attributes of Web-based health communities including ease of use and perceived synchronicity influence patients' intention to interact through convenience of Web-based health communities, whereas the attributes of physical health facilities such as inaccessibility and discontinuity affect patients' intention to interact through inconvenience of physical health facilities. We employed the survey method to validate our hypothesized relationships. Through developing the measurement instruments, we collected 334 valid answers from Web health community users and utilized partial least square to analyze the data.

Results: Ease of use (t_{311}=2.924, P=.004) and perceived synchronicity (t_{311}=2.353, P=.019) were found to influence convenience of Web-based health communities significantly, whereas inaccessibility (t_{311}=3.189, P=.002) and discontinuity (t_{311}=3.149, P=.002) were found to impact inconvenience of physical health facilities significantly. Meanwhile, both convenience of Web-based health communities (t_{311}=2.353, P=.019) and inconvenience of physical health facilities (t_{311}=2.787, P=.006) were found to affect patients' intention to interact with doctors in Web-based health communities significantly. Therefore, all the proposed hypotheses were supported.

Conclusions: Through including factors from both Web-based health communities and physical health facilities, we can understand patients' intention to interact comprehensively. This study not only contributes to literature of doctor-patient interaction and Web-based health platforms but also provides implications to promote doctor-patient interaction online and offline.

(J Med Internet Res 2019;21(6):e13693) doi:10.2196/13693

KEYWORDS

medical informatics; telemedicine; patients; physicians; community network; psychological theory; social theory

XSL•FO
RenderX

Introduction

Background

Patients in China are increasingly using internet to understand and treat their health issues. Almost 32.7% Chinese internet users have sought support online for their health issues, and the size of Chinese internet health care market is expected to reach 90 billion RMB [1]. Meanwhile, Chinese doctors spent more than 50% of their online time to do health-related activities [2]. Among all the internet applications, Web-based health communities provide the opportunities for interaction between patients and doctors in China [3].

Web-based health communities allow patients to establish discussion groups with doctors and other patients who have common interest in certain health problems and provide online tools for patients to contact doctors directly [4]. Patients can interact with any doctors they prefer in Web-based health communities. Therefore, communication mode between patients and doctors have been changed from single direction communication to multiple direction communication [5]. In the meantime, interactions between doctors and patients have been shown to promote patients' healthy behaviors and health information seeking, which could enhance their health status and alleviate the severity of many diseases including diabetes, hypertension, stroke, and Alzheimer disease [6,7]. Therefore, promoting interaction between patients and doctors in Web-based health communities is important for patients' health. Previous literature has studied the factors that influence interaction between patients and doctors. For example, Schillinger et al [8] found functional health literacy had positive association with quality of physician–patient communication among diabetes patients. Hsu et al [9] showed doctors' use of computers associated with patients' general satisfaction, communication about medical issues, and comprehension of medical decisions made. Wald et al [10] revealed that the use of internet could lead to more informed patients and in turn better health care choices. Roter et al [11] investigated the role of gender in physician-patient interaction by using meta-analysis method and found that female physicians engaged in communication with patients longer than male physicians. Schouten and Meeuwesen [12] showed doctors might behave differently toward different ethnic patients in doctor-patient communication. Although previous literature has studied some predictors of interaction between doctors and patients, few of them considers the context of Web-based health communities and related factors.

Meanwhile, considering the large population and unique health care system in China, the role of Web-based health communities in China may be different [13]. In addition, special cultural, social, and institutional forces in China may make patients' interaction with doctors in Web-based health communities distinct [14]. Thus, it is necessary to study patients' interaction with doctors in Web-based health communities in China context. Therefore, we propose our research questions as follows: *What are the predictors of patients' intention to interact with doctors in Web-based health communities in China?*

To address the research questions, we arrange the following sections as follows: we first apply the two-factor theory and service convenience theory to explore the predictors of patients' intention to interact with doctors and their underlying mechanisms. Afterwards, we establish our research model by developing corresponding hypotheses. Methods and Results follow. Discussion, implications, and future directions are in the last section.

Theoretical Foundation

Service convenience theory provides us the main theoretical perspective to understand patients' interaction and the basis to construct the working mechanisms of patients' interaction intention's predictors. Meanwhile, based on two-factor theory, we can decide the predictors of patients' intention to interact with doctors in Web-based health communities.

Service convenience theory assumes that consumers' perceived convenience is influenced by different characteristics of service, firm-related factors, and individual consumer differences, and it also determines their evaluation of services [15]. Characteristics of service are the main determinants of perceived convenience. Given both Web-based health communities and physical health facilities provide the place for doctors to provide services to patients, they may affect the quality of doctors' service. Therefore, it is possible to use service convenience theory to formulate the underlying mechanisms of predictors. To be specific, we propose convenience of Web-based health communities, and inconvenience of physical health facilities serve as the mechanisms for predictors of patients' interaction.

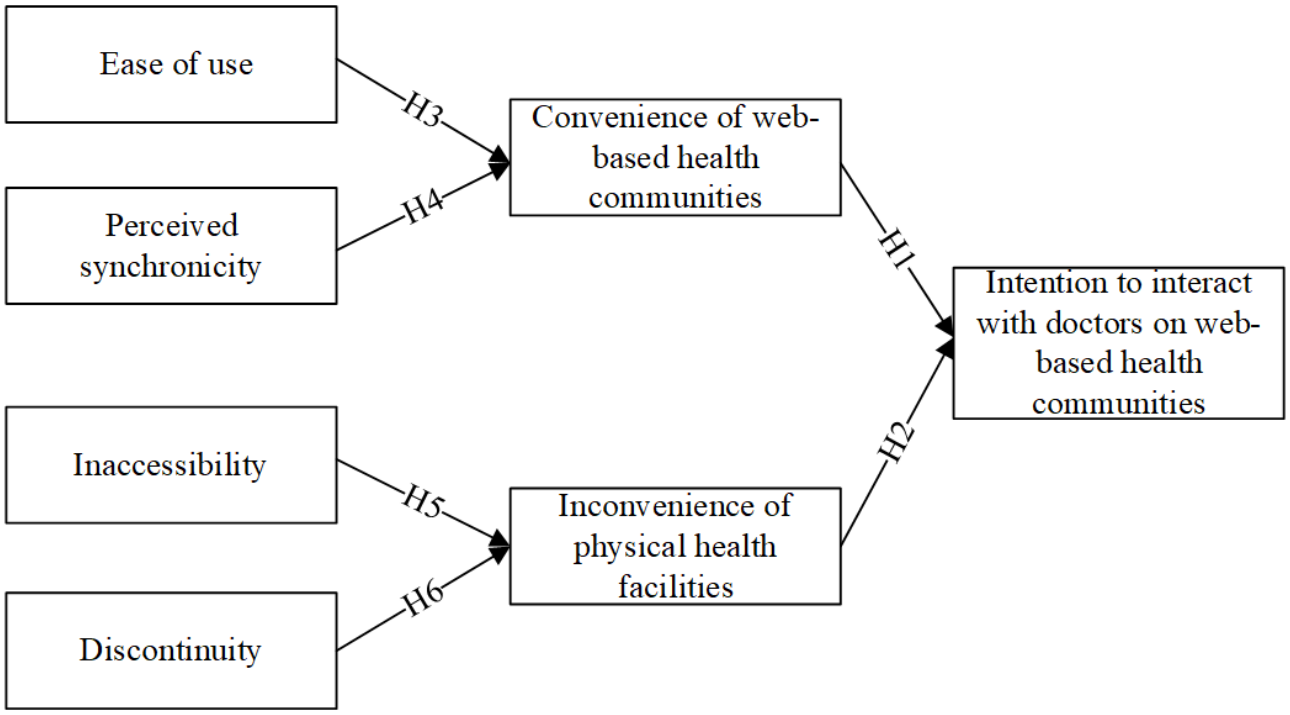
Two-factor theory is originally used to understand satisfaction in organization context [16]. This theory suggests that there are 2 categories of factors that influence satisfaction: hygiene and motivation factors. Hygiene factors may lead to dissatisfaction when not present, whereas motivation factors can lead to satisfaction when present. Therefore, convenience of Web-based health communities and inconvenience of physical health facilities could be understood by two-factor theory as convenience and inconvenience could correspond to hygiene and motivation factors. Toward motivation factors, the attributes of Web-based health communities could be the sources, whereas attributes of physical health facilities could link to the hygiene factors.

Research Model and Hypotheses Development

On the basis of service convenience theory, we propose that convenience of Web-based health communities and inconvenience of physical health facilities influence patients' intention to interact with doctors in Web-based health communities. Meanwhile, according to two-factor theory, we hypothesize that ease of use and perceived synchronicity affect convenience of Web-based health communities, whereas inaccessibility and discontinuity impact inconvenience of physical health facilities. The assumed relationships are depicted in Figure 1.

XSL•FO
RenderX

Figure 1. Research model and hypothesized relationships.



Convenience, Inconvenience, and Intention to Interact

Convenience can be defined as people's perception of time and effort involved in performing some tasks [15]. Therefore, patients feel convenient when they use less time and effort in using Web-based health communities, whereas patient perceive inconvenience when physical health facilities cost them too much time and effort. Convenience have been shown to facilitate the adoption and use of different information technologies [17,18]. As convenience of Web-based health communities can give value to patients through offering medical resources directly to patients and responding patients' health demands quickly [19], the value from Web-based health communities make patients adopt and use Web-based health communities. Meanwhile, inconvenience of physical health facilities may bring barriers or even threats to patients' health problems solving. Thus, the barriers or threats from inconvenience of physical health facilities may make patients avoid using physical health facilities and switch to Web-based health communities. On the basis of the above reasoning and explanations, we hypothesize that:

- H1: Convenience of Web-based health communities influences patients' intention to interact with doctors in Web-based health communities positively.
- H2: Inconvenience of physical health facilities influences patients' intention to interact with doctors in Web-based health communities positively.

Ease of Use, Perceived Synchronicity, and Convenience of Online Health Communities

Ease of use is the degree to which using Web-based health communities is free of effort [20]. Thus, ease of use of Web-based health communities may make patients cost little effort in using Web-based health communities and lead to patient feel using Web-based health communities be convenient. Previous literature has also shown the relationship between ease of use and convenience. For example, Yoon and Kim [18] showed that perceived ease of use positively influenced convenience of wireless LAN. Lai and Chang [21] found that perceived ease of use affected convenience of electronic book readers positively. Therefore, we can hypothesize that:

- H3: Ease of use influences convenience of Web-based health communities positively.

Perceived synchronicity is identified as 1 component of interactivity and defined as the degree to which the reception of messages is simultaneous to their sending in communication [22]. When patients can receive the feedback from doctors quickly in Web-based health communities no matter how far away they are, they will save their time and effort. Thus, patients may perceive convenience of using Web-based health communities. As such, we hypothesize that:

- H4: Perceived synchronicity influences convenience of Web-based health communities positively.

Inaccessibility, Discontinuity, and Inconvenience of Physical Health Facilities

Inaccessibility is the opposite of accessibility and refers to the degree to which patients cannot use health care service to achieve the best health outcomes [23]. Inaccessibility contains several dimensions including unaffordability, physical inaccessibility, unacceptability of services, and inadequacy of service supply [24]. Inaccessibility of health care services in physical health facilities becomes the barriers for patients to access the health care services they need and cost patients much time and effort in physical health facilities. Therefore, inaccessibility of health care services in physical health facilities may cause inconvenience for patients. Afterwards, we can hypothesize that:

XSL•FO
RenderX

- H5: Inaccessibility influences inconvenience of physical health facilities positively.

Discontinuity is the degree to which the whole health care process for patients is not coherent and connected [25]. It also includes some aspects such as informational discontinuity, management discontinuity, and relational discontinuity. Discontinuity may go against building trust and bond between doctors and patients, and afterwards the compliance toward doctors' medical treatment and advices [26]. Patients need to spend more time and effort to use discontinuous health care services. Thus, discontinuity of health care services in physical health facilities may also lead to inconvenience for patients. Afterwards, we can hypothesize that:

- H6: Discontinuity influences inconvenience of physical health facilities positively.

Methods

Measurement Instruments

To verify the proposed relationship in this study, survey method is employed in this study. Measurement instruments are developed based on previous literature. Items for intention to interact with doctors are adapted from Jang et al [27], items for convenience of Web-based health communities and inconvenience of physical health facilities are adapted from Yoon and Kim [18], items for ease of use are from Venkatesh and Davis [28], items for perceived synchronicity are adapted from Liu [29], and items for inaccessibility and discontinuity are adapted from Wang and Haggerty [30]. Besides the constructs studied in our research model, we also consider several control variables, which are mentioned in previous literature including age, gender, education, intensity and length of using Web-based health communities, and perceived health [7]. Meanwhile, the experience of filling online questionnaires of respondents is included in the questionnaire to test the non-naïve effect [31]. All items were measured by 5-point Likert scales, which are anchored from "1=strongly disagree" to "5=strongly agree."

Given our measurement instrument was adapted and developed from English literature, back translation method is used to translate it into Chinese. One of the bilingual authors first translated the English version of measurement instrument into Chinese version. Afterwards, another bilingual author back translated the Chinese version into English version. The 2 authors checked the consistency between the 2 English versions by comparing them. Inconsistency is resolved through discussion between the 2 authors. After the translation, we have checked the quality of measurement instrument preliminarily by conducting a pretest through interviewing 9 experts in the area of medical informatics, health management, and 16 users of Web-based health communities. Through the pretest, the experts and users provide comments and suggestions to improve the quality of our measurement instrument. On the basis of the comments and suggestions, we decide our measurement instrument and present it in Multimedia Appendix 1.

Data Collection

Given China has the largest population of internet users, we collected data from China [32]. To access to users of Web-based health communities efficiently, we employed the paid survey service from a leading online market research company. The service of market research firms allows for the efficiently administering online surveys and recruiting voluntary, motivated, and willing research participants for different research purposes [33]. This company has a sampling base that contained 2.6 million members, and more than 1 million of them were the daily questionnaire respondents. Given the objective of this study is to investigate patients' intention to interact with doctors in Web-based health communities, respondents with experience of interacting with doctors in Web-based health communities were randomly invited to fill out our questionnaires randomly from the sample base. The institutional review board of Tongji Medical College, Huazhong University of Science and Technology has approved our study procedures (No 2017S319). Through 3 weeks of deployment, we received a total of 427 responses.

To ensure the data quality and reduce social desirability bias, several actions are taken during the data collection. First, attention-trap and reverse-coded questions were used in the questionnaire to reduce single-method bias and check whether respondents were reading all questions fully and honestly. Second, several screening questions were set to check whether the respondents were patients who interacted with doctors in Web-based health communities such as whether the respondents had interacted with doctors in Web-based health communities, which Web-based health communities' respondents used most, and whether respondents were the members of Web-based health communities. Finally, the cases with missing values or similar values for all questions were discarded. Almost 93 answers were excluded, which leaves 334 valid responses. Therefore, the percent of invalid responses is 27.8% (93/334), which is reasonable [34].

Results

Demographic Information

The demographic information of our sample is showed in Table 1.

XSL•FO
RenderX

Table 1. Demographic information of sample participants.

Characteristics	n (%)
Age (years)	
<25	61 (18.3)
25-30	122 (36.5)
>30	151 (45.2)
Gender	
Male	138 (41.3)
Female	196 (58.7)
Education	
High school	7 (2.1)
College	294 (88)
Master's degree and above	33 (9.9)
Intensity of using Web-based health communities (hour/day)	
<0.5	274 (82)
0.5-1	54 (16.2)
>1	6 (1.8)
Length of using Web-based health communities (years)	
<1	126 (37.7)
1-5	164 (49.1)
>5	44 (13.2)

Reliability and Validity

To examine the reliability and validity of our measurement instrument, we conduct the confirmatory factor analysis by using SmartPLS 2.0.3M, which is a software developed by SmartPLS GmbH and is used for structural equation modeling [35]. Table 2 presents the results on reliability. The values of Cronbach alpha and composite reliabilities are all above .7; thus, confirming good reliability for all constructs [36]. Next, the values of average variance extracted (AVE) of each construct are all above 0.5, and loadings for each items are also all above 0.7; thus, reflecting good convergent validity [33]. Finally, Table 3 presents the results on validity. The values of square roots of AVEs are all greater than the inter-construct correlations, thus showing good discriminant validity [37]. Therefore, reliability and validity of our measurement instrument are acceptable.

Besides the tests of validity and reliability, common method bias was also considered. First, we went through the values of correlation coefficients among contracts in Table 3 to check whether they are too high (r>.90) and found all the values were not beyond the threshold. Second, we conducted Harman single factor test by the principle component analysis. A total of 7 factors were extracted, and the first extracted factor in the unrotated solution accounted for 23.53%, which is less than 50% [38]. Finally, we used marker variable technique to test the bias. Perceived organizational support was chosen as the marker variable as it was not relevant to our study [39]. The analysis result showed the average value of correlation coefficients between perceived organization support and other variables was only 0.156. Therefore, common method bias was not present in our study.

XSL•FO
RenderX

Table 2. Construct reliability and convergent validity.

Construct	Factor loadings	Composite reliability	Average variance extracted	Cronbach alpha
Intention to interact (ITI)				
ITI1	0.8363	0.836	0.6301	.7088
ITI2	0.7369	—[a]	—	—
ITI3	0.805	—	—	—
Convenience of Web-based health communities (CE)				
CE1	0.8641	0.8501	0.6556	.7388
CE2	0.7124	—	—	—
CE3	0.8443	—	—	—
Inconvenience of physical health facilities (IC)				
IC1	0.8189	0.8309	0.6211	.7013
IC2	0.7576	—	—	—
IC3	0.7866	—	—	—
Ease of use (EOU)				
EOU1	0.8101	0.8503	0.6543	.7376
EOU2	0.8021	—	—	—
EOU3	0.8145	—	—	—
Perceived synchronicity (PSY)				
PSY1	0.7957	0.8392	0.635	.7128
PSY2	0.8038	—	—	—
PSY3	0.791	—	—	—
Inaccessibility (IAY)				
IAY1	0.7462	0.844	0.5758	.7547
IAY2	0.7044	—	—	—
IAY3	0.8352	—	—	—
IAY4	0.7435	—	—	—
Discontinuity (DC)				
DC1	0.7442	0.834	0.5569	.7368
DC2	0.7848	—	—	—
DC3	0.7321	—	—	—
DC4	0.7225	—	—	—

[a]—Not applicable.

XSL•FO
RenderX

Table 3. Discriminant validity.

	ITI[a]	CE[b]	IC[c]	EOU[d]	PSY[e]	IAY[f]	DC[g]
ITI	0.7938[h]	—[i]	—	—	—	—	—
CE	0.3484	0.8097[h]	—	—	—	—	—
IC	0.3826	0.2979	0.7881[h]	—	—	—	—
EOU	0.3477	0.4166	0.2576	0.8089[h]	—	—	—
PSY	0.3192	0.3642	0.2057	0.3835	0.7969[h]	—	—
IAY	0.2415	0.289	0.4002	0.1852	0.0593	0.7588[h]	—
DC	0.1245	0.2369	0.4039	0.1788	0.0991	0.3697	0.7463[h]

[a]ITI: intention to interact.

[b]CE: convenience of Web-based health communities.

[c]IC: inconvenience of physical health facilitates.

[d]EOU: ease of use.

[e]PSY: perceived synchronicity.

[f]IAY: inaccessibility.

[g]DC: discontinuity.

[h]The square roots of average variances extracted.

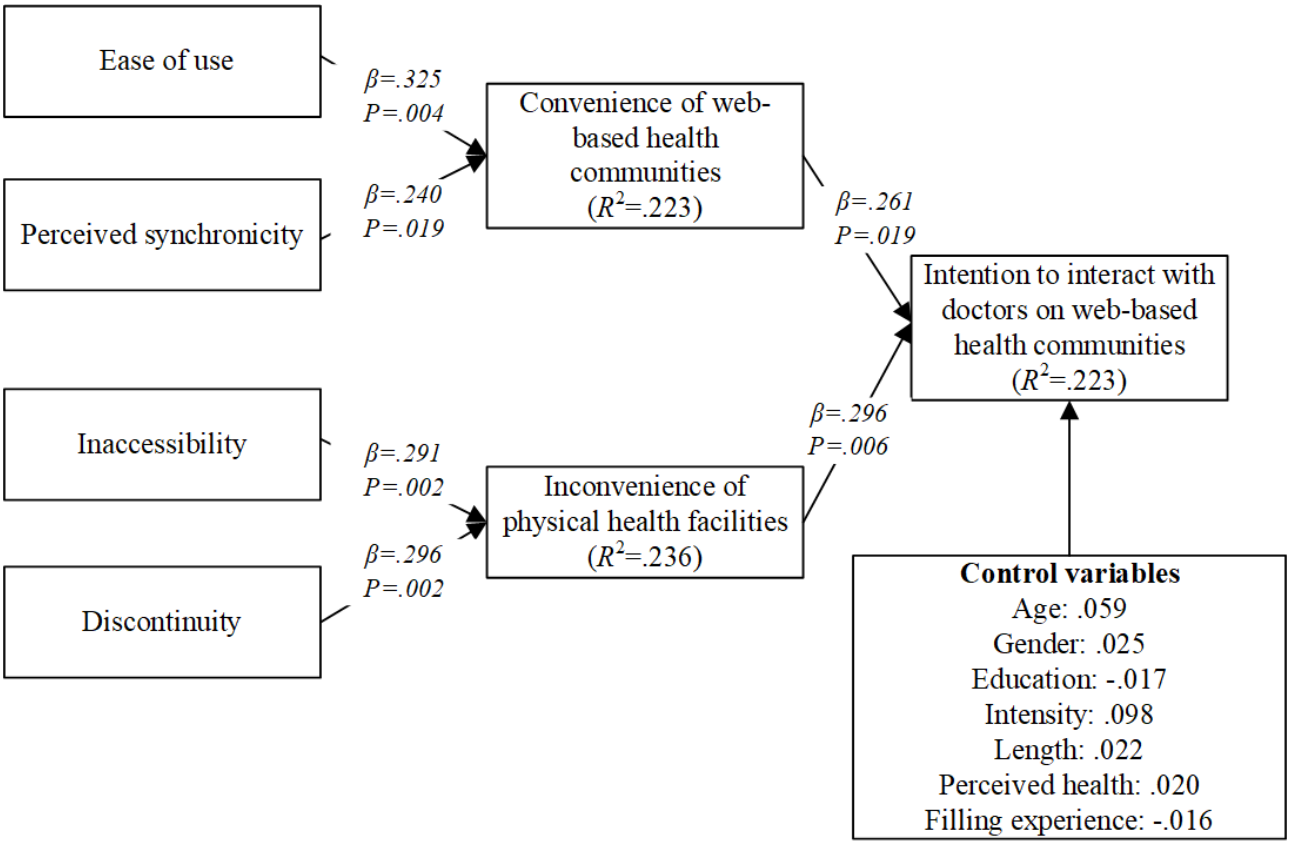
[i]—: not applicable.

Analysis Results of Structural Model

Through conducting the bootstrapping analysis in PLS, we test the hypothesized relationships in this study. The analysis results are shown in Figure 2. Both convenience of Web-based health communities and inconvenience of physical health facilities influence patients' intention to interact with doctors in Web-based health communities. Therefore, H1 and H2 are supported. These results confirm the effectiveness of service convenience theory. Meanwhile, ease of use and perceived synchronicity are both shown to influence convenience of Web-based health communities. Therefore, H3 and H4 are supported. These results suggest the attributes of Web-based health communities do determine patients' perception of Web-based health communities. Finally, inaccessibility and discontinuity are found to influence inconvenience of physical health facilities. Therefore, H5 and H6 are supported. There results reveal the attributes of physical health facilities also decide patients' feeling of physical health facilities.

XSL•FO
RenderX

Figure 2. Analysis results of hypothesized model.



Discussion

Principal Findings

This study explores the predictors of patients' intention to interact with doctors in Web-based health communities. On the basis of service convenience theory, we propose convenience of Web-based health communities and inconvenience of physical health facilities influence patients' intention to interact with doctors in Web-based health communities directly. Meanwhile, based on two-factor theory, we propose that ease of use and perceived synchronicity influence convenience of Web-based health communities, whereas inaccessibility and discontinuity influence inconvenience of physical health facilities. All the proposed relationships are manifested by using the survey method. These results imply that our research model can help understand patients' intention to interact with doctors in Web-based health communities adequately.

Implications

This study provides both theoretical and practical implications. For the theoretical implications, we first integrate service convenience theory and two-factor theory to understand patients' intention to interact with doctors in Web-based health communities. Through the integration, we can see patients' intention to interact through a complement theoretical view. Second, we contribute to patient-doctor interaction literature by studying their predictors from patients' view on Web health community context first. We consider the factors from Web-based health communities including ease of use and perceived synchronicity. Meanwhile, we also investigate the factors from physical health facilities such as inaccessibility and discontinuity. These factors answer our research questions directly. Finally, we explore the mechanisms of the predictors of patients' intention to interact with doctors in Web-based health communities. The mechanisms help us improve our theoretical modeling of patients' intention to interact.

Besides the theoretical implications, this study also provides its practical utility. First, this study confirms that Web-based health communities provide an important space for the interaction between doctors and patients. Second, there are implications for both health policy makers and Web health community managers. On the basis of factors from Web-based health communities, Web health community managers should provide enough tools and ensure the functionality of their communities for patients to communicate with doctors. Meanwhile, health policy makers should focus on the accessibility and continuity of physical health facilities, not just patients' satisfaction. Third, the measure scales of convenience and inconvenience could be the direct indicators to reflect patients' intentional or actual interaction with doctors in Web-based health communities. Finally, telehealth, which use the information communication technology to provide health care service distantly, have been developed in China for many years [40,41]. However, a recent study about telehealth in China has shown that the Chinese patients' adoption of telehealth is quite low although they are aware of telehealth [42]. Therefore, to explore the drivers of telehealth usage is necessary. As Web-based health communities can take some form of telehealth, the factors explored in our study can be examined to promote the utilization of telehealth.

XSL•FO
RenderX

Limitation and Future Direction

The limitations in this study can be the basis for future research directions. First, only 4 factors based on two-factor theory are included. On the basis of our current level of explained variance of convenience and inconvenience, more factors could be relevant. Second, although patients' intention could link to their actual behavior, it is necessarily to study patients' actual interaction with doctors in Web-based health communities as the actual behaviors can give more direct evidence of the effectiveness of our research model. Finally, we use cross-sectional data to study patients' intention to interact and did not account the dynamic of variables in our model. A longitudinal study would be useful as a follow-up.

Conclusions

This study examines the predictors of patients' intention to interact with doctors in Web-based health communities based on service convenience theory and two-factor theory. Through including factors from both Web-based health communities and physical health facilities, we establish a theoretical model to understand patients' intention to interact comprehensively. Through collecting data randomly from the sample pool of a research company, we validate the proposed research model. This study not only contributes to literatures of doctor-patient interaction and we health communities but also provides implications to promote doctor-patients interaction online and offline.

Acknowledgments

This work was partially supported by grants from the National Natural Science Foundation of China (No 71801100, 71671073, 71602063, 71704059, 71701059 and 71810107003), and the Fundamental Research Funds for the Central Universities (No 2018KFYYXJJ063 and 2019kfyXJJS172).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Measurement instrument.

[PDF File (Adobe PDF File), 64KB - jmir_v21i6e13693_app1.pdf]

References

1. Li N. [Billion Euro_Technology and Industry Innovation Service Platform]. 2018. [90 billion yuan market size! The Internet medical industry has entered an expansion phase] URL: https://www.iyiou.com/p/71492.html [accessed 2019-02-13] [WebCite Cache ID 769AZlvF7]
2. Li A. Kantar China Insights. 2017. Chinese doctors increasingly prefer WeChat over apps URL: https://cn-en.kantar.com/business/health/2017/chinese-doctors-increasingly-prefer-wechat-over-apps/ [accessed 2019-04-18] [WebCite Cache ID 77kBlZcI6]
3. Libert B, Beck M. Forbes. 2017. Digital Platforms Are The Key To Improving Your Health In 2017 URL: https://www.forbes.com/sites/barrylibert/2017/01/04/digital-platforms-are-the-key-to-improving-health/#37728f693b0f [accessed 2018-08-20] [WebCite Cache ID 769B3wEvb]
4. Maloney-Krichmar D, Preece J. A multilevel analysis of sociability, usability, and community dynamics in an online health community. ACM Trans Comput-Hum Interact 2005 Jun 1;12(2):201-232. [doi: 10.1145/1067860.1067864]
5. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471 [FREE Full text] [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
6. Wu T, Deng Z, Zhang D, Buchanan PR, Zha D, Wang R. Seeking and using intention of health information from doctors in social media: the effect of doctor-consumer interaction. Int J Med Inform 2018 Jul;115:106-113. [doi: 10.1016/j.ijmedinf.2018.04.009] [Medline: 29779712]
7. Wu T, Deng Z, Feng Z, Gaskin DJ, Zhang D, Wang R. The effect of doctor-consumer interaction on social media on consumers' health behaviors: cross-sectional study. J Med Internet Res 2018 Feb 28;20(2):e73 [FREE Full text] [doi: 10.2196/jmir.9003] [Medline: 29490892]
8. Schillinger D, Bindman A, Wang F, Stewart A, Piette J. Functional health literacy and the quality of physician-patient communication among diabetes patients. Patient Educ Couns 2004 Mar;52(3):315-323. [doi: 10.1016/S0738-3991(03)00107-1] [Medline: 14998602]
9. Hsu J, Huang J, Fung V, Robertson N, Jimison H, Frankel R. Health information technology and physician-patient interactions: impact of computers on communication during outpatient primary care visits. J Am Med Inform Assoc 2005;12(4):474-480 [FREE Full text] [doi: 10.1197/jamia.M1741] [Medline: 15802484]
10. Wald HS, Dube CE, Anthony DC. Untangling the Web--the impact of internet use on health care and the physician-patient relationship. Patient Educ Couns 2007 Nov;68(3):218-224. [doi: 10.1016/j.pec.2007.05.016] [Medline: 17920226]

XSL•FO
RenderX

11. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: a meta-analytic review. J Am Med Assoc 2002 Aug 14;288(6):756-764. [Medline: 12169083]
12. Schouten BC, Meeuwesen L. Cultural differences in medical communication: a review of the literature. Patient Educ Couns 2006 Dec;64(1-3):21-34. [doi: 10.1016/j.pec.2005.11.014] [Medline: 16427760]
13. Yip W, Subramanian SV, Mitchell AD, Lee DT, Wang J, Kawachi I. Does social capital enhance health and well-being? Evidence from rural China. Soc Sci Med 2007 Jan;64(1):35-49. [doi: 10.1016/j.socscimed.2006.08.027] [Medline: 17029692]
14. Li L, Gao P, Mao J. Research on IT in China: a call for greater contextualization. J Inform Technol 2014 Sep;29(3):208-222. [doi: 10.1057/jit.2014.15]
15. Berry LL, Seiders K, Grewal D. Understanding service convenience. J Mark 2018 Oct 10;66(3):1-17. [doi: 10.1509/jmkg.66.3.1.18505]
16. Maddox RN. Two-factor theory and consumer satisfaction: replication and extension. J Consum Res 1981 Jun;8(1):97. [doi: 10.1086/208845]
17. Eastin MS. Diffusion of e-commerce: an analysis of the adoption of four e-commerce activities. Telem Inform 2002 Aug;19(3):251-267. [doi: 10.1016/S0736-5853(01)00005-3]
18. Yoon C, Kim S. Convenience and TAM in a ubiquitous computing environment: the case of wireless LAN. Electron Commer Res Appl 2007 Mar;6(1):102-112. [doi: 10.1016/j.elerap.2006.06.009]
19. Lin K, Lu H. Predicting mobile social network acceptance based on mobile value and social influence. Internet Res 2015 Feb 2;25(1):107-130. [doi: 10.1108/IntR-01-2014-0018]
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
21. Lai J, Chang C. User attitudes toward dedicated e‐book readers for reading. Online Inform Rev 2011 Aug 9;35(4):558-580. [doi: 10.1108/14684521111161936]
22. Liu Y, Shrum LJ. What is interactivity and is it always such a good thing? Implications of definition, person, and situation for the influence of interactivity on advertising effectiveness. J Advert 2002 Dec;31(4):53-64. [doi: 10.1080/00913367.2002.10673685]
23. Institute of Medicine (US) Committee on Monitoring Access to Personal Health Care Service. In: Millman M, editor. Access to Health Care in America. Washington, DC: National Academies Press; 1993.
24. Gulliford M, Figueroa-Munoz J, Morgan M, Hughes D, Gibson B, Beech R, et al. What does 'access to health care' mean? J Health Serv Res Policy 2002 Jul;7(3):186-188. [Medline: 12171751]
25. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R. Continuity of care: a multidisciplinary review. Br Med J 2003 Nov 22;327(7425):1219-1221 [FREE Full text] [doi: 10.1136/bmj.327.7425.1219] [Medline: 14630762]
26. Sudhakar-Krishnan V, Rudolf MC. How important is continuity of care? Arch Dis Child 2007 May;92(5):381-383 [FREE Full text] [doi: 10.1136/adc.2006.099853] [Medline: 17449521]
27. Jang H, Olfman L, Ko I, Koh J, Kim K. The influence of on-line brand community characteristics on community commitment and brand loyalty. Int J Electron Commer 2014 Dec 8;12(3):57-80. [doi: 10.2753/JEC1086-4415120304]
28. Venkatesh V, Davis F. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000:186-204 [FREE Full text] [doi: 10.1287/mnsc.46.2.186.11926]
29. Liu Y. Developing a scale to measure the interactivity of websites. J Advert Res 2003 Jun 1;43(2):207-216. [doi: 10.2501/JAR-43-2-207-216]
30. Wong ST, Haggerty J. University of British Columbia. 2013. Measuring patient experiences in primary health care: a review and classification of items and scales used in publicly-available questionnaires URL: https://open.library.ubc.ca/cIRcle/collections/facultyresearchandpublications/52383/items/1.0048528 [accessed 2019-02-11] [WebCite Cache ID 769C0oghG]
31. Goodman J, Paolacci G. Crowdsourcing consumer research. J Consum Res 2017;44(1):196-210. [doi: 10.1093/jcr/ucx047]
32. Internet World Stats. 2018. Top 20 countries with the highest number of internet users URL: https://www.internetworldstats.com/top20.htm [accessed 2019-02-13] [WebCite Cache ID 769DHW9k2]
33. Lowry PB, D'Arcy J, Hammer B, Moody GD. "Cargo Cult" science in traditional organization and information systems survey research: a case for using nontraditional methods of data collection, including Mechanical Turk and online panels. J Strateg Inform Syst 2016 Oct;25(3):232-240. [doi: 10.1016/j.jsis.2016.06.002]
34. Tavares J, Oliveira T. Electronic health record patient portal adoption by health care consumers: an acceptance model and survey. J Med Internet Res 2016 Mar 2;18(3):e49 [FREE Full text] [doi: 10.2196/jmir.5069] [Medline: 26935646]
35. Ringle CM, Wende S, Will A. SmartPLS 2. URL: http://www.smartpls.de [accessed 2019-05-27] [WebCite Cache ID 78gCTTUab]
36. Nunnally J, Bernstein I. Psychometric Theory. New York: McGraw-Hill; 1967.
37. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39. [doi: 10.2307/3151312]
38. Podsakoff PM, Organ DW. Self-reports in organizational research: problems and prospects. J Manage 2016 Sep 4;12(4):531-544. [doi: 10.1177/014920638601200408]
39. Lindell MK, Whitney DJ. Accounting for common method variance in cross-sectional research designs. J Appl Psychol 2001;86(1):114-121. [doi: 10.1037/0021-9010.86.1.114]

XSL•FO
RenderX

40. Perednia D, Allen A. Telemedicine technology and clinical applications. J Am Med Assoc 1995;273(6):483-488. [doi: 10.1001/jama.273.6.483]
41. Scott R, Mars M. Telehealth in the developing world: current status and future prospects. Smart Homecare Technol TeleHealth 2015 Feb:25. [doi: 10.2147/SHTT.S75184]
42. Chen P, Xiao L, Gou Z, Xiang L, Zhang X, Feng P. Telehealth attitudes and use among medical professionals, medical students and patients in China: a cross-sectional survey. Int J Med Inform 2017 Dec;108:13-21. [doi: 10.1016/j.ijmedinf.2017.09.009] [Medline: 29132618]

Abbreviations

AVE: average variance extracted

Edited by G Eysenbach; submitted 18.02.19; peer-reviewed by PCI Pang, C Khalil; comments to author 28.03.19; revised version received 08.05.19; accepted 10.05.19; published 07.06.19.

Please cite as:
Wu T, Deng Z, Chen Z, Zhang D, Wang R, Wu X
Predictors of Patients' Intention to Interact With Doctors in Web-Based Health Communities in China: Cross-Sectional Study
J Med Internet Res 2019;21(6):e13693
URL: http://www.jmir.org/2019/6/e13693/
doi:10.2196/13693
PMID:31199296

©Tailai Wu, Zhaohua Deng, Zhuo Chen, Donglan Zhang, Ruoxi Wang, Xiang Wu. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mining of Textual Health Information from Reddit: Analysis of Chronic Diseases With Extracted Entities and Their Relations

Vasiliki Foufi[1,2*], PhD; Tatsawan Timakum[3*], BA, MA; Christophe Gaudet-Blavignac[1,2*], BSc CS, MMed; Christian Lovis[1,2], MD, MPH, FACMI; Min Song[3], PhD

[1]Division of Medical Information Sciences, University Hospitals of Geneva, Geneva, Switzerland

[2]Faculty of Medicine, University of Geneva, Geneva, Switzerland

[3]Department of Library and Information Science, Yonsei University, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Min Song, PhD
Department of Library and Information Science
Yonsei University
50 Yonsei-ro, Seodaemun-gu
Seoul, 120-749
Republic of Korea
Phone: 82 22123 2405
Fax: 82 2393 8348
Email: min.song@yonsei.ac.kr

Abstract

Background: Social media platforms constitute a rich data source for natural language processing tasks such as named entity recognition, relation extraction, and sentiment analysis. In particular, social media platforms about health provide a different insight into patient's experiences with diseases and treatment than those found in the scientific literature.

Objective: This paper aimed to report a study of entities related to chronic diseases and their relation in user-generated text posts. The major focus of our research is the study of biomedical entities found in health social media platforms and their relations and the way people suffering from chronic diseases express themselves.

Methods: We collected a corpus of 17,624 text posts from disease-specific subreddits of the social news and discussion website Reddit. For entity and relation extraction from this corpus, we employed the PKDE4J tool developed by Song et al (2015). PKDE4J is a text mining system that integrates dictionary-based entity extraction and rule-based relation extraction in a highly flexible and extensible framework.

Results: Using PKDE4J, we extracted 2 types of entities and relations: biomedical entities and relations and subject-predicate-object entity relations. In total, 82,138 entities and 30,341 relation pairs were extracted from the Reddit dataset. The most highly mentioned entities were those related to oncological disease (2884 occurrences of cancer) and asthma (2180 occurrences). The relation pair anatomy-disease was the most frequent (5550 occurrences), the highest frequent entities in this pair being cancer and lymph. The manual validation of the extracted entities showed a very good performance of the system at the entity extraction task (3682/5151, 71.48% extracted entities were correctly labeled).

Conclusions: This study showed that people are eager to share their personal experience with chronic diseases on social media platforms despite possible privacy and security issues. The results reported in this paper are promising and demonstrate the need for more in-depth studies on the way patients with chronic diseases express themselves on social media platforms.

(J Med Internet Res 2019;21(6):e12876) doi:10.2196/12876

KEYWORDS

social media; chronic disease; data mining

XSL•FO
RenderX

Introduction

Background

People are often concerned about their health status and a range of medical issues, especially when it comes to complex or chronic diseases that can take a long time to treat or monitor. Patients often desire easy access to information about diseases and symptoms to understand their condition and to facilitate self-management of diseases without total reliance upon interaction with a physician [1]. Patients with chronic diseases in particular use social media to seek and provide social, emotional, and practical support [2]. Therefore, social media information can influence patients' decisions to manage their chronic condition [3].

Social media platforms may support patients in their search for medical products or provide suggestions to promote healthy behavior and can improve health education as they allow people to write about their experiences with diseases, drugs, symptoms, and treatments. In recent years, social media platforms have grown quickly, with the public, patients, and health professionals sharing their experiences, looking for information, and interacting with others.

Currently, more than 74% of internet users connect to social media, and 42% of the internet users take advantage of social media for health information. Moreover, 32% of social media users in the United States share about their health care experiences and family's struggle stories and 29% search for health information via social media platforms to observe other patients' experiences with their diseases [3]. Furthermore, 51% of those who live with a chronic disease have used the internet for information about health topics such as details of a specific disease, medical procedures, drugs, medical devices, or health insurances [4].

With its growing number of users, social media has become a powerful tool that can promote information sharing about health care, provide feedback from users, and foster support systems [5]. In addition, the existence of social media platforms enables researchers to learn and discover the health experiences and feeling of patients and potentially discover new knowledge in health science. For example, user conversation content from health-related online forums, such as blogs, Twitter, and Facebook, has already been analyzed to find the clusters of breast cancer symptom [6], examine smoking [7], and understand the user discourse and describe social media interactions about obesity prevention [8]. In particular, Reddit has been used as a data source for similar studies [9-12].

The interactions between individuals on social media and the information they share constitute an important new source of data that can be used, on one hand, to understand the impact of drugs, diseases, and medical treatments on patients outside controlled clinical settings and, on the other hand, to comprehend health-related behavior.

Discovering public knowledge in social media text constitutes a challenge for researchers and health care providers. To achieve this goal, various text mining approaches, such as topic modeling, information extraction, and visualization, exist.

Biomedical Entity and Relation Extraction

In the era of biomedical text mining, bioentities and their relations have arisen as a challenge to discover new knowledge. To mine the huge amounts of unstructured data, automatic information extraction tools have been conceived and developed based on several approaches. There are multiple systems developed for the identification and analysis of relations between diseases, drugs, and genes, such as Extraction of Drugs, Genes and Relations, a natural language system that extracts information about drugs and genes relevant to cancer from the biomedical literature [13]. Extraction of drug-disease treatment pairs from the published literature was also carried out [14,15]. To extract health social media information, adverse drug reactions and drug indications from a Spanish health forum were examined [16] using MeaningCloud [17], a multilingual text analysis engine based on a distant-supervision method to detect relations between drugs and side effects and used them to classify the relation instances.

PKDE4J2.0 is a system that extracts bioentities and their relations with the aim to discover biomedical scientific knowledge. It is based on a dictionary to automatically tag bioentities according to their types and a set of predefined rules used for relation extraction. PKDE4J2.0 can be applied for knowledge search, knowledge network construction, and knowledge inference [18]. PKDE4J1.1 was used to investigate drug-disease interactions in article abstracts from PubMed Central for making drug-symptom-disease triples [19]. This tool was also applied in biomedical literature to extract biomedical verbs to present a relation type between 2 entities [20] and on full-text papers to extract biological entities from diseases and genes and construct a knowledge network [21].

Health Information Extraction From Social Media Platforms

A large number of patients, caregivers, and health professionals use social media platforms to discuss mental health issues. They also constitute an important data source for researchers. Machine learning and statistical methods were used to discriminate online messages between depression and control communities using mood, psycholinguistic processes, and content topics extracted from the posts generated by members of these communities [22]. Users are interested in searching for treatment-related information, communicating with physicians to share their feelings about treatment effectiveness and side effects, discussing questions in health communities, and gaining knowledge about their conditions [23]. User-generated content from these platforms contains valuable information [24]. Their posts reflect what users think and feel about their medical experiences and often attract the attention of other patients, caregivers, and doctors.

Lu et al [25] mined data from online health communities and used text clustering integrating medical domain–specific knowledge to investigate patient needs and interests. Their results show that compared with existing methods, the addition of medical domain–specific features into their feature sets achieved significantly better clustering than was achieved without the addition of those features. Moreover, there were significant differences in hot topics on different kinds of disease

XSL•FO
RenderX

discussion platforms. Health-related posts on social media were analyzed to investigate the polarity of opinions online, performing sentiment analysis [26]. Medical terms, including those related to conditions, symptoms, treatments, effectiveness, and side effects, were extracted to generate a virtual document addressing each question raised by members of the community. Then latent Dirichlet allocation (LDA) was modified by adding a weighting scheme known as conditional LDA to cluster virtual documents with similar distributions of medical terms into a conditional topic (C-topic). Finally, the clustered C-topics were analyzed according to sentiment polarities and physiological and psychological sentiments. Identification of topics of patients' discussions on (1) Facebook about breast cancer and (2) cancerdusein.org was performed [27]. These topics were assigned to functional and symptomatic dimensions by applying LDA topic modeling and identified relations between the topics and the questionnaires.

Among others, Denecke [1] reported that "user-generated content on the web has become a new source of useful information to be added to the conventional methods of collecting clinical data."

In terms of biomedical information extraction, previous studies relied on formal research and individual case studies to identify biomedical information. These approaches include observations of changes in patients [28], meta-analysis of data from relevant databases [29], and surveys of cancer patients [30]. However, the scientific literature is generally limited to subscribers, and electronic medical records are not publicly available for reasons of patient privacy [31]. Moreover, these sources do not provide a complete understanding of how patients suffering from a chronic disease feel and how they express these feelings.

Using data from conversations between patients on social media platforms provides valuable information for researchers, physicians, and health care providers. This data source is different from, and complementary to, that obtained from conventional experimental methods.

Research Objectives

Therefore, a social media platform (Reddit) was chosen as the data source for this research that aimed to answer the following questions:

1. Which biomedical entities are prominent in the health social media platforms?
2. What types of entities are related in the corpus?
3. How do people express themselves about chronic diseases on social media platforms?

Methods

Data Collection

The data used for this research were extracted from disease-specific subreddits of the social news and discussion website Reddit [32]. Forums such as Reddit tend to have sharp contrast when compared with similar offline groups; for instance, people are likely to discuss problems that they do not feel comfortable to discuss face to face [33]. As of 2013, Reddit's official statistics included 56 billion page views, 731 million unique visitors, 40,855,032 posts, and 404,603,286 comments [34]. In particular, the subreddit about cancer numbers 22,429 subscribers and 75 posts per day [35]. These numbers demonstrate the external validity of Reddit. Another reason for having chosen Reddit as a data source is that the language of text posts is more structured than in other social media platforms such as Twitter.

Reddit's core functionality is the sharing of text-based posts with others who may or may not be members of the site. The subforum function allows the creation of designated spaces for users to congregate and interact with each other over a shared interest. Those subforums are called *subreddits*. A finite set of 19 subreddits related to chronic diseases was empirically selected for analysis. The choice of the specific subreddits was based on medical expertise and on the impact of these diseases on the quality of everyday life of patients.

As the main goal was the detection of relations between entities and of the way people suffering from chronic diseases express themselves in social media and not the study of characteristics of specific chronic diseases, the posts from the 19 subreddits were merged in a single dataset.

All of these subreddits host public content. In this research, no populational study has been performed. The study focuses on the expression of feelings and not on the identity of people sharing their experiences. From each post, only the title of the post and the body or textual content was extracted without additional information related to their authors.

The study was submitted to the Swiss Ethical Committee who concluded to a decision of nonconsideration provided that the collected data are not identifiable.

Lexicosemantic Resources

Lexicosemantic resources were constructed and incorporated into the tool. These resources included a list of stop words and biomedical dictionaries of diseases, drugs, anatomy, procedures, symptoms, side effects, and findings created from clinical health care terminologies such as the Systematized Nomenclature of Human and Veterinary Medicine - Clinical Terms [36], the National Library of Medicine's controlled vocabulary thesaurus [37], the Gene Ontology knowledgebase [38], the Kyoto Encyclopedia of Genes and Genomes database [39], and the DrugBank database [40]. Semantic relations properties were attributed to 4558 biomedical verbs extracted from the Unified Medical Language System [41].

The dictionaries were enriched with lemmas extracted from the corpus; for instance, *chemo*, *AML* (acute myeloid leukemia), *take care*, *support*, and *fight*.

Description of the Tool

In this research, the PKDE4J version 2.0 tool [42] was used. This text mining system consists of 2 modules: entity extraction and relation extraction.

Entity Extraction Module

This module integrates dictionary-based entity extraction and rule-based relation extraction into a highly flexible and extensible framework. The Stanford CoreNLP pipeline [43]

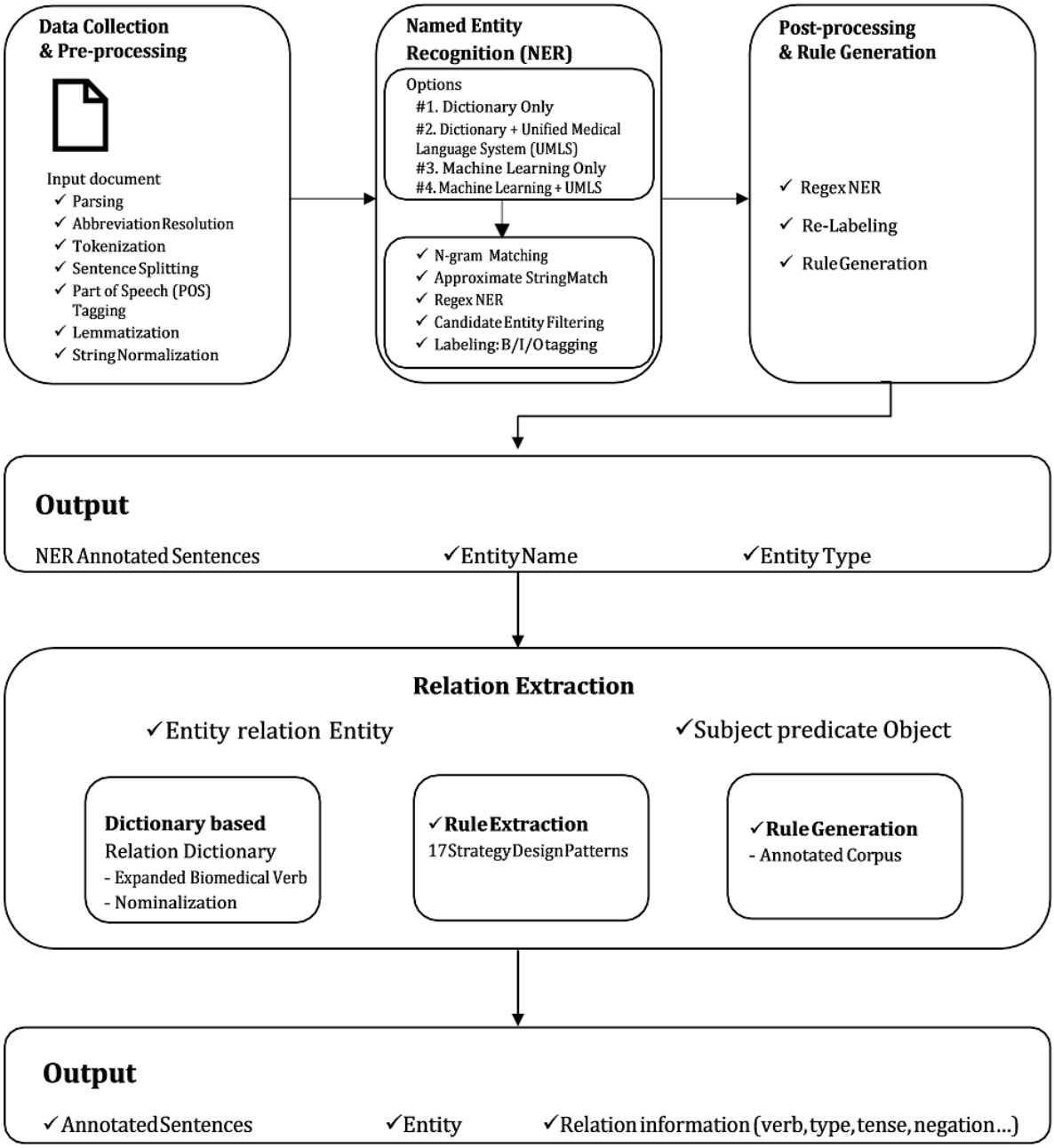
XSL•FO
RenderX

was modified to make it suitable for advanced dictionary-based entity extraction. The entity extraction module consists of 4 major submodules: preprocessing, dictionary loading, entity annotation, and postprocessing. PKDE4J can analyze entities and relations from both structured and unstructured text.

Relation Extraction Module

The relation extraction workflow identifies directed qualified relations starting from sentences from which 2 or more entities have been extracted by the entity extraction module. The relation extraction module takes a list of verbs and nominalization terms that are employed to identify relations of interest. After extracting entities from a sentence, further relation extraction algorithms are executed to construct rules for the extraction of relations of entities. A set of 20 dependency parsing–based rules is at the core of the relation extraction module and provides an ontologically enriched structure for sentences by annotating edges with dependency types. To extract relations, the system identifies a verb, which may be located between entities and contains relational characteristics, then, it checks the bioverb list to determine the relation between the entities (Figure 1) [44].

Figure 1. The workflow of the PKDE4J text mining system.



XSL•FO
RenderX

Visualization

The Gephi platform [45] was used to visualize the network of chronic diseases in the corpus. To build a graph, the *k*- shortest paths routing algorithm was applied. The graph visualization tool was then used to map the chronic disease entities. A PageRank of terms was computed to rank the important entities in the network; therefore, entities ranked highly by PageRank have the highest impact.

Validation of Entity Extraction

To evaluate the performance of the tool on entity extraction, 1000 posts randomly selected from the entire corpus were manually validated. The entities were evaluated as correct or incorrect based on the following specific guidelines.

Findings and Symptoms

This category refers to a phenomenon that is experienced by a person or described by a clinician and cannot be considered as a disease in the context, for example, "This news makes me feel anxiety."

Disease Names

This category refers to an abnormal condition of a human, animal, or plant that causes discomfort or dysfunction [46]. As also mentioned in the previous category, the context helps to distinguish between a disease and a symptom or finding. For example, in the sentence "After trying which dosage is good, my insomnia is thankfully gone again," *insomnia* refers to a disease, whereas in the sentence "I have had symptoms of insomnia within the last months," *insomnia* describes a symptom/finding.

Side Effects

This category includes a symptom/finding or a disease that is caused by a treatment in the context. For example:

> *Since beginning treatment have woken with bouts of nausea...*

Procedure

Procedure refers to any intervention carried on someone related to physical mental or social health. For example:

> *...treatment which would include surgery and radiation/chemotherapy according to his oncologist*

Results

Data Collection

A dataset of 17,624 text posts was semiautomatically collected using crawlers accessing public streams. Table 1 shows the subreddits used for this research, the number of posts per subreddit, and the proportion of corpus representation of each subreddit:

Table 1. Sources used for the data collection.

Subreddit name	Number of posts	Proportion of posts from each subreddit in the corpus
r/cancer	5210	26.9
r/MultipleSclerosis	1902	9.8
r/rheumatoid	1783	9.2
r/CrohnsDisease	1722	8.9
r/Asthma	1600	8.3
r/testicularcancer	1384	7.1
r/Parkinsons	1042	5.4
r/Hashimotos	1022	5.3
r/Alzheimers	927	4.8
r/breastcancer	794	4.1
r/braincancer	623	3.2
r/pancreaticcancer	397	2.1
r/lymphoma	387	2.0
r/leukemia	223	1.2
r/kidney	107	0.6
r/multiplemyeloma	104	0.5
r/thyroidcancer	63	0.3
r/lungcancer	41	0.2
r/skincancer	15	0.1
Total	19,346	100

XSL•FO
RenderX

After sorting the data corpus, duplicate posts and those with no relevant meaning, such as advertising posts and posts containing only a hyperlink, were removed. The final corpus comprises 17,580 posts (2,137,115 tokens).

Biomedical Entity and Relation Extraction

The PKDE4J system performed named entity extraction and 2 types of relation extraction: relations between biomedical entities and between subjects, predicates, and objects on the sentence level. The system's output is a corpus annotated with entities and information about their relation.

The entities are either simple terms or complex structures referring to diseases, anatomy, procedures, findings, symptoms, side effects, or drugs. In total, PKDE4J extracted 82,138 entities from the Reddit dataset, as shown in Table 2. The entity names and entity types were allocated to the 7 categories of the biomedical dictionaries. The 10 most frequent entity names followed by the number of occurrences in the corpus are displayed in Table 3. It should be noted that the terms are given in the text in the form found in the corpus. Therefore, abbreviated terms have not been expanded.

As displayed in the table, 29,669 disease entities were extracted representing 1341 unique diseases; 19,956 anatomy entities, of which 369 are distinct anatomical terms; 11,549 procedures of 296 different types; 6256 symptoms entities describing 65 symptoms; 5351 entities representing side effects of 321 different types; and 35 different drug names (616 in total). The most highly represented diseases are oncological (*cancer*, *breast cancer*, *tumor*, *leukemia*, and *lymphoma*) or relate to asthma. The anatomy category contains a range of anatomical terms. Specifically, *blood* is the most frequent term. Other widely used anatomical terms are *back*, *brain*, *hand*, *hair*, *breast*, *chest*, *heart*, and *neck*. The procedures category comprises terms referring to chemical treatments (*chemo*), surgery, laboratory test (*blood test*), social interventions (*advice* and *listening*), and others.

The most frequent symptom mentioned in the corpus is *pain* (472 occurrences). *Fatigue*, *inflammation*, *nausea*, and *cough* are some of the symptoms commonly reported by patients or relatives in the dataset. In the side-effect category, the most frequent entities are *anxiety*, *stress*, *swelling*, *crying*, and *fear* followed by *disability* and *worry*. The most commonly reported drug is *prednisone* followed by *morphine*, *salbutamol*, and *tramadol*.

Validation of Entity Extraction

Among the 5151 extracted entities, 3682 were correctly labeled by the system, whereas 1469 were attributed with incorrect labels. The performance of the system was 71.48%.

Next, an error analysis was performed on the incorrectly labeled entities. Errors were classified into 3 categories:

1. Lexical errors (488/1469, 33.21%): the term *breast* is an anatomical term, but in the post, the compound term *breast cancer* appears. However, the system failed to extract the entire entity.
2. Dictionary errors (550/1469, 37.44%), for example, *air* and *aspergillus* were falsely listed as an anatomical term and as a drug name, respectively.
3. Ambiguous concepts (431/1469, 29.33%): the term *bleeding* could be either a disease name or a symptom.

Table 2. Entity extraction results.

Entity types	Diseases	Anatomy	Procedures	Findings	Symptoms	Side effects	Drugs
Extracted entities, n	29,669	19,956	11,549	8741	6256	5351	616
Entity names, n	1341	369	296	483	65	321	35

Table 3. Ten most frequent entities by type.

Diseases	Anatomy	Procedures	Findings	Symptoms	Side effects	Drugs
Cancer (2884)	Blood (1542)	Chemo (1914)	Related (521)	Pain (2648)	Anxiety (683)	Prednisone (417)
Asthma (2180)	Back (1034)	Treatment (1909)	Lump (359)	Fatigue (639)	Stress (373)	Morphine (33)
All (2163)	Brain (962)	Surgery (1909)	Suffering (333)	Inflammation (472)	Swelling (348)	Salbutamol (33)
Breast cancer (804)	Hand (656)	Advice (774)	Confused (305)	Scared (273)	Crying (245)	Tramadol (26)
Can (745)	Head (627)	Radiation (627)	Problem (304)	Nausea (244)	Mass (220)	MRSA[a] (14)
Tumor (631)	Hair (549)	Biopsy (366)	Attack (277)	Hurt (205)	Fear (215)	Aspirin (11)
Disease (563)	Breast (535)	Chemotherapy (338)	Energy (270)	Sore (202)	Disability (164)	Omeprazole (9)
TSH[b] (506)	Chest (511)	Blood test (268)	Terrified (266)	Numb (195)	Worry (142)	Seretide (6)
Depression (414)	Heart (503)	Infusion (199)	Tired (251)	Cutting (104)	Fall (129)	Citrus (5)
Lymphoma (348)	Neck (459)	Listening (158)	Follow up (249)	Tingling (101)	Discomfort (120)	Echinacea (5)

[a]MRSA: methicillin-resistant *Staphylococcus aureus*.

[b]TSH: thyroid stimulating hormone.

XSL•FO
RenderX

Relation Extraction

The system extracted 2 entities (entity 1 and entity 2) found in the same sentence and linked with a relation and then it attributed the type of entities. For instance, entity 1, *Borderline* (disease) co-occurs with *High blood pressure* (symptom) in the sentence. In total, 30,341 relation pairs were extracted, as shown in Table 4.

Of the 30,341 relation pairs, the most frequent entity relation pairs and their number of co-occurrences are shown in Table 5.

The relations between anatomy and disease entity types are the most frequent (5550 pairs). The pair disease-disease co-occurs 4668 times, and the pair anatomy-anatomy appears 3595 times.

Table 6 contains the 5 most frequent entities per relation pair.

Table 4. Example of entity relation extraction.

Analysis result	Entity 1	Entity 1 type	Entity 2	Entity 2 type	Sentence from post
Output 1	Borderline	disease	High blood pressure	symptom	Prior to that, I was fat mid forties male borderline high HDL, high blood pressure but ZERO issues with thyroid or immune issues.
Output 2	Optic neuritis	disease	Multiple sclerosis	symptom	She said that I have something called optic neuritis and that about half the time people get it and they don't know why but the other half its because someone has multiple sclerosis.
Output 3	Syndrome	disease	Nerve	anatomy	I had bilateral optic neuritis significantly worse in my left eye in Late August September and I was also simultaneously diagnosed with Browns Syndrome which they're not 100% convinced on as it may have been misdiagnosed 6th nerve palsy.

Table 5. Most frequent entities per relation pair.

Relation pair	Co-occurrences, n
anatomy-anatomy	3595
anatomy-disease, disease-anatomy	5550
anatomy-procedure, procedure-anatomy	1730
anatomy-symptom, symptom-anatomy	1227
anatomy-side effect, side effect-anatomy	1081
disease-disease	4668
disease-procedure, procedure-disease	2540
disease-finding, finding-disease	2128
disease-side effect, side effect-disease	1502
disease-symptom, symptom-disease	1080
finding-finding	303
finding-anatomy, anatomy-finding	1362
procedure-procedure	1023
procedure-finding, finding-procedure	430
side effect-side effect	256

Table 6. The 5 most frequent entities per relation pair.

Rank	Pair of entity 1 and entity 2					
	Anatomy|Anatomy	Disease|Disease	Disease|Side effect	Disease|Anatomy	Disease|Procedure	Disease|Finding
1	Back|Hair	Cancer|ALL	Depression|Anxiety	Cancer|Lymph	Cancer|Surgery	Asthma|Attack
2	Neck|Lymph	Asthma|Allergy	Asthma|Anxiety	Tumor|Blood	Tumor|Surgery	Cancer|Suffering
3	Brain|Lungs	Tumor|Seminoma	Cancer|Fear	Asthma|Lungs	Breast cancer|Treatment	ALL|Follow up
4	Lungs|Lymph	ALL|Asthma	ALL|Swelling	ALL|Blood	Asthma|Advice	Depression|Suffering
5	Head|Hair	Depression|Fatigue	Aches|Anxiety	TSH|Blood	ALL|Treatment	Exercise|Muscle tension

XSL•FO
RenderX

The most frequent entities in the pair disease-anatomy is *Cancer|Lymph*, in the pair disease-disease is *Cancer|ALL*, and in the pair anatomy-anatomy is *Back|Hair*.

To summarize, once the most frequent entities were extracted, the results were processed according to the shortest path between each entity pair to produce the graph shown in Figure 2. Among 2561 nodes and 13,405 edges, this entity network shows that *pain* highly co-occurs with other entities in the network (biggest node, weighted at 0.022461), followed by *cancer* (PageRank score at 0.018057) and *surgery* (PageRank score at 0.015443).

The node *pain* has connections with other nodes, including *fatigue*, *inflammation*, *stomach*, *joints*, *cancer*, and *chemo*. The node *cancer* is strongly linked to *chemo*, *surgery*, *treatment*, *ALL (acute lymphoblastic leukemia)*, *anxiety*, and *blood*. The nodes of *surgery*, *chemo*, and *treatment* are linked to diseases and body parts. Finally, the entity nodes relating to mental health, such as *anxiety* and *depression*, also appear in the network and associate with other bioentity types. Table 7 shows the most frequent entities and the corresponding PageRank scores:

Figure 2. Biomedical entity network.

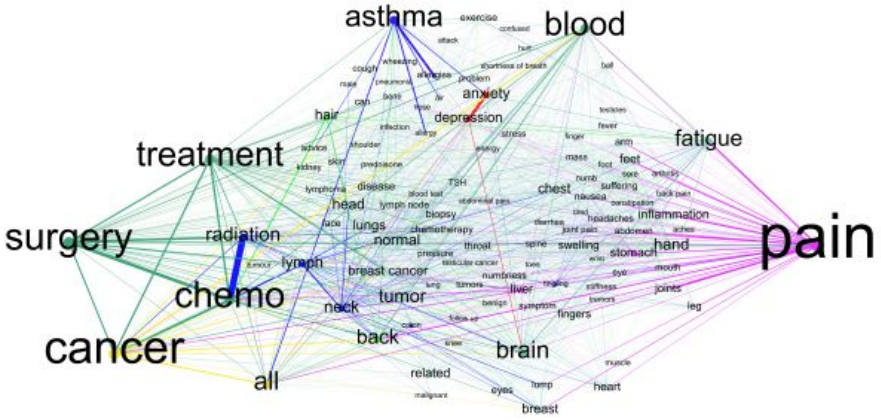


Table 7. The most frequent entities and the corresponding PageRank scores.

Label	PageRank score
Pain	0.022461
Cancer	0.018057
Surgery	0.015443
Chemo	0.014954
Treatment	0.014275
Blood	0.013841
Asthma	0.012554
All	0.010931
Brain	0.010311
Fatigue	0.009626
Back	0.008574
Radiation	0.008317
Tumor	0.007814
Neck	0.006968
Hand	0.006618
Lymph	0.006442
Normal	0.006176
Anxiety	0.006108
Head	0.005933
Hair	0.005762

XSL•FO
RenderX

Subject-Predicate-Object Entity Relation Extraction

The system extracted 69,263 subject or object entities. The top 10 entities are shown in Table 8. In total, 41,068 relations were extracted and the results were classified into 2 types of subjects: subject pronoun (*I*, *you*, *he*, *she*, *it*, *we*, and *they*) and subject noun (*treatment*). The relation pairs are divided into 19,645 pairs of subject pronoun-object entities and 21,423 pairs of subject noun-object entities.

Table 9 shows 2 examples of the subject-predicate-object relation extraction: the subject (for example *I*, *he*, *anyone*, *it*, and *asthma*), the predicate (verbs such as *have*, *get*, and *increase*), the object (terms such as *eczema*, *allergies*, *childhood asthma*, *my cough*, and *allergy shots*), and the sentence of the corresponding post.

Subject Pronoun-Predicate-Object

The subject pronoun-predicate-object relation extraction demonstrates that the most frequent subject pronoun is *I* (11,691 times, including *I've*, *I'm*, and *I'd*). Some examples are shown in Table 10.

Subject Noun-Predicate-Object

Table 11 shows some examples of subject noun-predicate-object relation extraction. Among the 21,423 relation pairs, the most frequent subject nouns are diseases such as *asthma* (272 occurrences), including phrases such as *asthma anxiety*, *asthma attacks*, *my asthma*, *my asthma and allergies*, *my asthma flare*, and *cancer* (226 occurrences).

Table 8. The top 10 occurrences of subjects and objects.

Entity	Count, n
I	10,314
It	2832
Pain	2341
She	1501
He	1323
Cancer	1060
Asthma	1015
They	900
Me	719
You	597

Table 9. Examples of subject-predicate-object relation extraction results.

Analysis result	Subject entity	Predicate	Object entity	Sentence
Output 1	I	Had	the skin allergy test	I had the skin allergy test done and it came back positive for almost every kind of pollen and mold, etc.
Output 2	my blue inhaler	Increases	my asthma	I noticed consistently my blue inhaler increases my asthma about 30% after using it and I believe was the cause of a recent very bad asthma attack.

Table 10. Example of subject pronoun-predicate-object relation extraction.

Subject	Predicate	Object	Sentence
I	take	the typical seretide	I take the typical seretide morning and night and ventolin when l need it.
I	have	a deep and painful cough	I have a deep and painful cough that's been leaving me with back, chest, and side pains.
You	ever take	allergy shots	Hey, to you asthmatics who have allergy induced asthma, did you ever take allergy shots.
He	was given	Prednisone	He was given prednisone for that as well.
She	has	Asthma	My sister, who lives with me, started complaining to me about it, saying that she doesn't want me doing that when her daughter is home because she has asthma and it smokes up the house.

XSL•FO
RenderX

Table 11. Example of subject noun-predicate-object relation extraction.

Subject	Predicate	Object	Sentence
Asthma	is becoming way more than just	a physical issue	Asthma is becoming way more than just a physical issue, it's taking a toll on my mental health.
Cancer	had spread to	her bones	The doctor told her that cancer had spread to her bones and that she'll have to have injections for it?
Fever	is indeed mentioned as	a side effect	Ive also been using Modulair Montelukast Sodium, and fever is indeed mentioned as a side effect on my leaflet.
Hives	are from	allergies	They’re trying to tell me they are panic attacks but as far as I know hives are from allergies and they sometimes happen during my asthma attacks.
The depression	is occurring simultaneously with	the increased asthma symptoms	I noticed the depression is occurring simultaneously with the increased asthma symptoms and was wondering if there is a correlation and if anyone else has experienced this.
My milk allergy	was causing	my asthma	My milk allergy was causing my asthma.
My second course of prednisone	has been great for stopping	the wheezing	And I'm on my second course of prednisone which has been great for stopping the wheezing - even the rescue inhaler didn't help before.
The doctor	ruled out	pneumonia	Anyway, the doctor ruled out pneumonia and said I had caught a cold on the plane and it had triggered an asthma exacerbation.

Figure 3 demonstrates the most used subject and object entities. The results show that the most frequent subject is the pronoun *I* (PageRank score: 0.100619). The pronoun *It*, *She*, *He*, *They*, and *You* are also frequently used. Diseases, body parts, treatments, and symptoms are widely used as the subjects and/or objects as well as the possessive pronoun *my* (*my mother*, *my eyes*, and *my dad*). Table 12 presents the most frequent subject and object entities and the corresponding PageRank scores.

Social Media Language

Expressions that constitute specific terms developed on social media, such as *pm* (private message), *FWIW* (For What it's Worth) were identified in the corpus. “A common feature of microblog texts is the use of symbols in posts, such as the love-heart dingbat symbol” [47]. Emoticons such as  , and text-based emoticons such as *LOL* (laughing out loud), :), :-(, =), and :(are also frequent.

In addition, the corpus contains informal phrases such as “Rooting for you!” and “I’m still chugging along”; adjectives such as *loopy*, *drippy*, *dicey*, and *zonked*; and verbs such as *puke* that substitute for their equivalents in standard language.

Entities found in the Reddit corpus present numerous morphosyntactic variants. For example, the term *chemotherapy* was rarely found, but the short form *chemo* was frequently used. The disease name *Hodgkin’s Lymphoma* is as *Hodgkin Lymphoma*, *Hodgkins Lymphoma*, *Hodgkin disease*, and *HL*. Similarly, the entity name *Mixed Cellularity Classical Hodgkin Lymphoma* is found as *Mixed Cellularity Hodgkin Lymphoma*, *Mixed Cellularity Hodgkins Lymphoma*, and *MCCHL*. Moreover, there are many abbreviated forms of entity names, such as *ALL*, *AML*, *BRCA2* (breast cancer type 2), *CLL* (chronic lymphocytic leukemia), *CML* (chronic myeloid leukemia), *COPD* (chronic obstructive pulmonary disease), *DCIS* (ductal carcinoma in situ), and *GERD* (gastroesophageal reflux disease). When these forms were included in the disease dictionary, the system managed to detect them. Some examples are presented in Table 13.

XSL•FO
RenderX

Figure 3. Subject and object entity network.

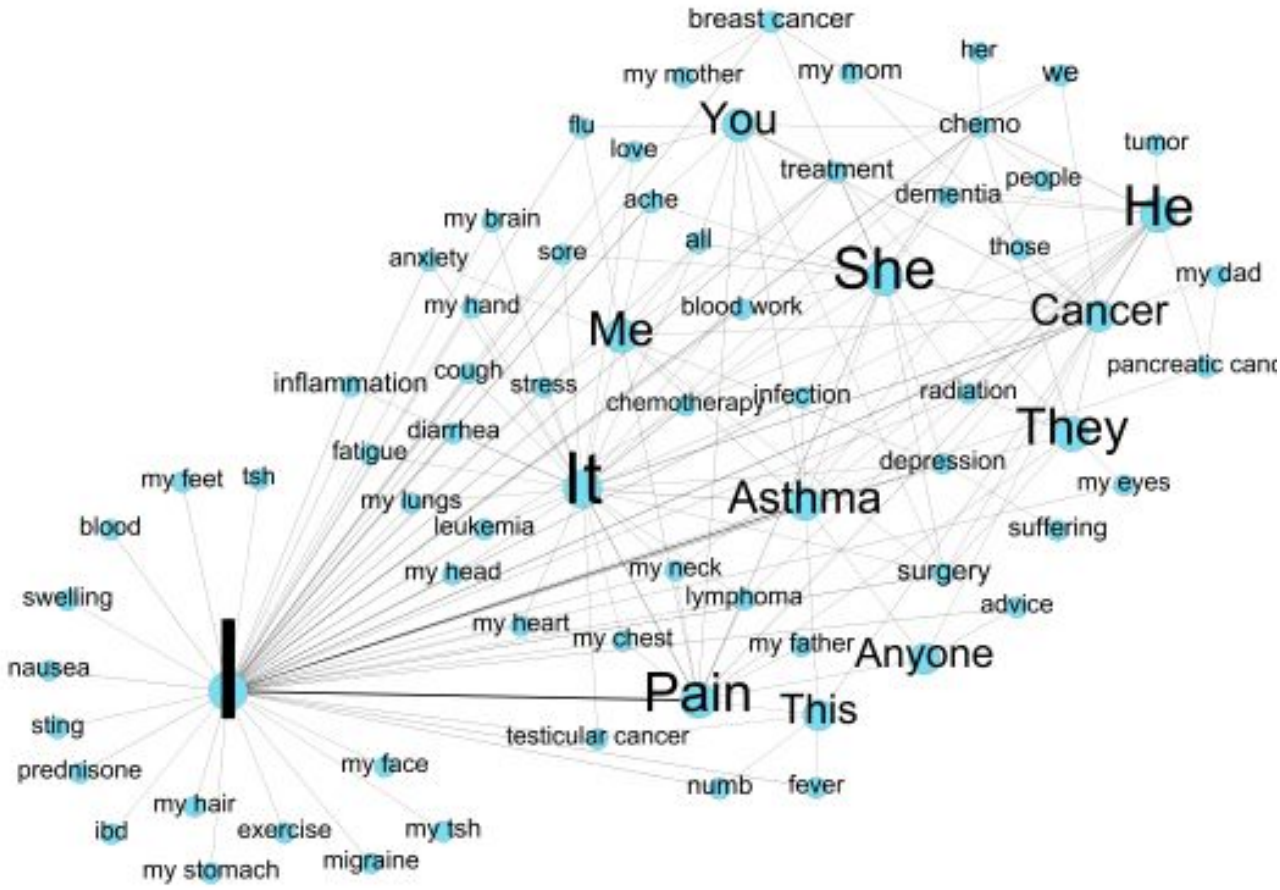


Table 12. The most frequent entities and the corresponding PageRank scores.

Label	PageRank score
I	0.100619
It	0.027234
Pain	0.020651
She	0.014206
He	0.012875
Cancer	0.00939
Asthma	0.009324
They	0.008927
Me	0.006934
You	0.005852

XSL•FO
RenderX

Table 13. Examples of disease entities in their abbreviated form.

Abbreviated form of entity name	Full entity name	Example from the corpus
ALL	Acute lymphoblastic leukemia	My boyfriend was diagnosed with ALL 2 years ago and stayed in remission after a few rounds of chemo.
AML	Acute myeloid leukemia	Diagnosed with AML this past Sept.
BRCA2	Breast cancer type 2	Her sister, my aunt, was diagnosed with breast cancer at 27 and was dead by 33 she tested positive for BRCA2 as well.
CLL	Chronic lymphocytic leukemia	My CLL is more of SLL, which is the same thing but presented in my lymph nodes.
CML	Chronic myeloid leukemia	25 years old, diagnosed with CML when I was 15.
COPD	Chronic obstructive pulmonary disease	Hes been smoking for over 40 years, has COPD and isnt in the greatest health generally overweight, inactive, etc.
DCIS	Ductal carcinoma in situ	We found out last week she has both DCIS and Invasive DCIS.
GERD	Gastroesophageal reflux disease	Sleep apnea can also worsen GERD, and GERD is known to worsen asthma.

Discussion

Principal Findings

In this paper, we collected user-generated chronic disease–related data from Reddit and extracted information pertinent to biomedical entities and their relations to examine the characteristics of the language used by users in this social media platform. Initially, the corpus was created by semiautomatically extracting posts from specific subforums of Reddit. Next, lexicosemantic resources from various sources were created. To perform the information extraction tasks—entity extraction and relation extraction—the PKDE4J text mining system was used. The system extracted 82,138 biomedical entities and 30,341 relations. These results indicate that the corpus contains a large amount of information.

Performance of the Tool

As described in the Results section, the system achieved a high performance in the named entity extraction task and the attribution of entity types (3682/5151, 71.48% extracted entities were correctly labeled). As already mentioned, the language used in Reddit is structured enough, with a satisfactory number of full sentences so the system managed to extract entities and their relations. The error analysis showed that the system failed to detect a number of entities or falsely attributed the entity type, because of lexical errors, to dictionaries' errors or to ambiguous concepts.

Entity Extraction

Entities prominent in the corpus refer to diseases, anatomical terms, procedures, findings, and symptoms. While interpreting the entities extracted from the corpus, it must be taken into account that the corpus was constructed by selecting subreddits created to share information about specific diseases. Therefore, it is expected that entities related to these diseases are the most likely to be represented. For instance, parts of the body affected by specific cancers, such as *breast* or *blood*, occur very frequently.

The most frequent disease entities in this corpus are oncologic diseases such as *cancer*, *ALL*, and *breast cancer*. Frequently mentioned nononcologic diseases are *asthma*, *depression*, and the generic entity *disease*. The entity *thyroid stimulating hormone* (TSH) is frequently mentioned, but it should be further classified in findings.

The most frequent anatomy entity is *blood*. This is explained primarily because of the numerous posts speaking about *leukemia* and *lymphoma*. Moreover, people often report the results of blood tests, a situation that increases the number of entities identified.

Terms tagged as *procedures* extracted from the corpus are mainly linked to oncologic diseases. About 2000 occurrences of *chemo* and *chemotherapy* were extracted. *Chemotherapy* is a significant procedure with numerous side effects. The fact that patients mention it at a high frequency shows that it is a treatment with a strong impact on quality of life and raises a lot of questions and worries for the patients involved. Social intervention procedures such as *listening* and *advice* are also frequent (see Table 3). This observation indicates that apart from technical information about treatment and surgeries, people also speak about the support they got during their disease or search for it in the community.

In medicine, symptoms can be difficult to differentiate from findings. This difference often resides in the context of the phenomenon. In the corpus, entities belonging to those categories as well as the side-effects category can be analyzed together to gain a better understanding of the results. Most frequent entities from these categories are closely related to the patient experiences and feelings. Concepts related to the feeling of fear are the most frequently present in this merged category: 7 out of 30 entities express feelings of fear or related with fear using the words *anxiety*, *stress*, *confused*, *scared*, *terrified*, *fear*, and *worry*. This is coherent with studies on cancer survivors that state the fear of cancer recurrence as almost universal among this population [48]. It appears that people with chronic conditions use social media to share feelings they have experienced. The chronic diseases selected in this corpus frequently imply severe impact on lifestyle and decrease life expectancy. Therefore, it is logical that *fear* and *anxiety* are prominent entities in the corpus.

XSL•FO
RenderX

Health-related quality of life in chronically ill patients is a known field in medical research since numerous years. Questionnaires such as the European Organisation for Research and Treatment of Cancer Quality of Life-C15-Palliative [49] or, more recently, the Functional Assessment of Cancer Therapy-General 7 [50] and Patient-Reported Outcomes Measurement Information System [51] are used routinely to assess it in those populations. When looking at the top concerns raised by patients suffering from cancer [50], it is interesting to note that they are in line with the top entities extracted from the corpus. More specifically, the most frequent nondisease entity extracted, *pain*, is a key item in multiple quality-of-life assessment questionnaires. This shows that the experiences that the patients share on social media platforms are coherent with what has been proven to have an impact on their life.

Overall, entities extracted from the corpus are coherent with similar studies conducted on health-related social media [27] and with validated evaluation of the quality of life of patients suffering from chronic diseases.

Relation Extraction

The relation extraction performed on this corpus shows that the most highly represented relation type identified is the *disease-anatomy* relation (5550 occurrences). The pairs most frequently representing a disease and its localization are *cancer-lymph*, *asthma-lungs*, and *tumor-blood*. This suggests that people using social media platforms to speak about their chronic diseases are willing to explain which disease they suffer from as well as the location of the disease. This propensity is probably linked to the fact that such subreddits are used to share life experiences and to find people with similar backgrounds. This commonality can be reassuring and informative for a person suffering from the same disease. To find these people and knowledge, it is valuable to share the nature of the disease and its anatomical location.

The second most frequent entity pair is *disease-disease* (4668 occurrences). Pairs of entities such as *cancer-ALL* (acute lymphoblastic leukemia) and *asthma-allergy* are frequent. This co-occurrence of diseases might be related to the fact that chronic diseases often lead to complications and to other diseases. For example, *asthma-allergy* was perceived from the sentence “have allergy induced asthma.”

The third most frequent entity pair is *anatomy-anatomy* (3595 occurrences). When looking at specific occurrences of this pair, the pairs *neck-lymph*, *brain-lungs*, and *lungs-lymph* are frequent. Another entity pair, *head-hair* is related to people speaking about the side effects of chemotherapy.

Relations linking *diseases* to *procedures* are present at a high frequency in the corpus (2540 occurrences). When looking specifically in this category, it is clear that the most frequently identified disease in the corpus, *cancer*, is also most highly represented in those relations.

The relations extracted from the corpus demonstrate that patients with chronic diseases are willing to share detailed information about their health condition in a structured manner, describing thoroughly the disease, its location, the symptoms it caused, and the effect of treatment.

Subject-Predicate-Object Entity Relation Extraction

The language patterns of subject-predicate-object relations demonstrate important characteristics of health social media language. As is apparent in the outputs, subject pronouns and object pronouns were frequently mentioned and were used mostly in the singular first-person pronoun, such as *I*, *me*, and *my*. These patterns are related to the way individuals share personal or family experiences (“I-had-a bad cold or sinus infection,” “Allergens-explains-my severe asthma,” “It-is making-my heartburn,” and “Anyone-develop-eczema”) and feelings (“I have a history of testicular cancer in family so Im pretty scared bht im hoping its nothing”). Also, patients or relatives, after having described their problem, treatment, and possible effects, often ask for advice, as shown in the following sentence:

> *I noticed the depression is occurring simultaneously with the increased asthma symptoms and was wondering if there is a correlation and if anyone else has experienced this?*

Social Media Language

Data derived from clinical narratives and research papers differ significantly from social media content. The language and style used by the authors as well as the content are different. From a linguistic point of view, medical blogs usually consist of syntactically correct sentences but can contain verbless clauses or sentences without subjects [52]. Abbreviations, enumerations, and citations of conversations, medical terms, and opinion-related words are used frequently in medical blog posts and websites. As stated in the study by Korkontzelos et al [53], “in social media, users rarely use technical terms.” Moreover, emoticons are very often used to convey emotion or to give contextual information to correctly understand a message (such as irony or sarcasm). The corpus processed in this research confirmed these observations.

Limitations

There are 2 major limitations of the PKDE4J tool with regard to the objectives of this paper. First, PKDE4J was initially developed for the processing of well-structured biomedical texts and not for social media text. This issue has a relatively less impact on this paper given that the entity extraction task is based on dictionaries. However, for tasks such as part-of-speech and sentence parsing needed for the extraction of relations, the informality of social media text poses a challenge. Second, the lack of terms from the dictionaries as well as lexical and semantic ambiguities lowered the performance of the system. For instance, abbreviations and acronyms can have multiple interpretations, and this can lead to ambiguities. In the current version of the system, these types of ambiguities are not handled. Consequently, all occurrences of *ALL* found in the corpus were extracted, even those that do not refer to the disease *Acute Lymphocytic Leukemia*. Also, the lexical unit *back* has sometimes been falsely recognized as a body part.

Conclusions

Data from social media platforms devoted to health can provide valuable information about the experiences of the patients involved. In this paper, we reported the application of an

XSL•FO
RenderX

information extraction approach using the PKDE4J tool to detect, extract, and visualize chronic disease entities and relations and to identify characteristics of the social media language in a corpus collected from Reddit.

In the Results section, we showed which disease entities are frequently mentioned and which are the most frequent relation pairs. Relation extraction demonstrated that the most frequent relation pair is the *disease-anatomy* pair and the subject-object relation pattern in the social media language is the use of the first-person pronoun provided that people share personal experiences.

Although data privacy and information sharing is becoming a major concern in research and legal frameworks, such as the General Data Protection Regulation law, have begun to set boundaries for the storage and sharing of information generated by users, it is interesting that despite those concerns, users are willing to share private health information in open social networks.

Further research should focus on the enrichment of dictionaries and adaptation of rules to common usages of social media language and the processing of emoticons for the sentiment analysis task. Finally, the identification of the type of semantic relations and the evaluation on the relation extraction results should be performed to assess the performance of the system in this task.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2018S1A3A2075114) and the University of Geneva, Switzerland.

Authors' Contributions

VF, TT, and CGB contributed equally to the manuscript.

Conflicts of Interest

CL is editor-in-chief for JMIR Medical Informatics.

References

1. Denecke K. Health Web Science: Social Media Data for Healthcare. New York: Springer International Publishing; 2015.
2. Patel R, Chang T, Greysen SR, Chopra V. Social media use in chronic disease: a systematic review and novel taxonomy. Am J Med 2015 Dec;128(12):1335-1350. [doi: 10.1016/j.amjmed.2015.06.015] [Medline: 26159633]
3. ReferralMD. 2017. 30 Facts & Stats on Social Media and Healthcare URL: https://getreferralmd.com/2017/01/30-facts-statistics-on-social-media-and-healthcare/ [accessed 2019-06-03] [WebCite Cache ID 78qwn2Dif]
4. Pew Research Center. Chronic Disease and the Internet URL: https://www.pewinternet.org/2010/03/24/chronic-disease-and-the-internet/ [accessed 2019-06-03] [WebCite Cache ID 78qx81xnX]
5. Moorhead S, Hazlett D, Harrison L, Carroll J, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23:e85. [doi: 10.2196/jmir.1933]
6. Marshall SA, Yang CC, Ping Q, Zhao M, Avis NE, Ip EH. Symptom clusters in women with breast cancer: an analysis of data from social media and a research study. Qual Life Res 2016 Mar;25(3):547-557 [FREE Full text] [doi: 10.1007/s11136-015-1156-7] [Medline: 26476836]
7. 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]
8. Chou WS, Prestin A, Kunath S. Obesity in social media: a mixed methods analysis. Transl Behav Med 2014 Sep;4(3):314-323 [FREE Full text] [doi: 10.1007/s13142-014-0256-1] [Medline: 25264470]
9. Sharma R, Wigginton B, Meurk C, Ford P, Gartner CE. Motivations and limitations associated with vaping among people with mental illness: a qualitative analysis of Reddit discussions. Int J Environ Res Public Health 2016 Dec 22;14(1):7 [FREE Full text] [doi: 10.3390/ijerph14010007] [Medline: 28025516]
10. Park A, Conway M. Tracking health related discussions on reddit for public health applications. AMIA Annu Symp Proc 2017;2017:1362-1371. [Medline: 29854205]
11. Pandrekar S, Chen X, Gopalkrishna G, Srivastava A, Saltz M, Saltz J, et al. Social media based analysis of opioid epidemic using Reddit. AMIA Annu Symp Proc 2018;2018:867-876 [FREE Full text] [Medline: 30815129]
12. Sumner SA, Galik S, Mathieu J, Ward M, Kiley T, Bartholow B, et al. Temporal and geographic patterns of social media posts about an emerging suicide game. J Adolesc Health 2019 Feb 25:1-7 (forthcoming). [doi: 10.1016/j.jadohealth.2018.12.025] [Medline: 30819581]
13. Rindflesch TC, Tanabe L, Weinstein JN, Hunter L. EDGAR: extraction of drugs, genes and relations from the biomedical literature. Pac Symp Biocomput 2000:517-528 [FREE Full text] [Medline: 10902199]
14. Xu R, Wang Q. Large-scale extraction of accurate drug-disease treatment pairs from biomedical literature for drug repurposing. BMC Bioinformatics 2013 Jun 6;14(1):181. [doi: 10.1186/1471-2105-14-181]

XSL•FO
RenderX

15. Zhu Y, Song M, Yan E. Identifying Liver Cancer and Its Relations with Diseases, Drugs, and Genes: A Literature-Based Approach. PLoS One 2016 May 19;11(5):e0156091 [FREE Full text] [doi: 10.1371/journal.pone.0156091] [Medline: 27195695]
16. Segura-Bedmar I, Martínez P, Revert R, Moreno-Schneider J. Exploring Spanish health social media for detecting drug effects. BMC Med Inform Decis Mak 2015 Jun 15;15(2):S6 [FREE Full text] [doi: 10.1186/1472-6947-15-S2-S6] [Medline: 26100267]
17. MeaningCloud. Text Analytics – MeaningCloud text mining solutions URL: https://www.meaningcloud.com/ [accessed 2019-06-03] [WebCite Cache ID 78r622IFx]
18. Song M, Kim M, Kang K, Kim YH, Jeon S. Application of public knowledge discovery tool (PKDE4J) to represent biomedical scientific knowledge. Front Res Metr Anal 2018:1-16 [FREE Full text] [doi: 10.3389/frma.2018.00007]
19. Song M, Kang K, An JY. Investigating drug–disease interactions in drug–symptom–disease triples via citation relations. J Assoc Inf Sci Technol 2018 Jul 30:1355-1368 [FREE Full text] [doi: 10.1002/asi.24060]
20. Kim Y, Beak S, Song M. Constructing Linguistic Verb Source for Relation Extraction. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. 2016 Presented at: CIKM'16; October 24-28, 2016; Indianapolis, Indiana, USA p. 2511-2512 URL: http://dtmbio.net/dtmbio2016/pdf/8.pdf
21. Amplayo R, Song M. ACL Member Portal. 2016. Building Content-driven Entity Networks for Scarce Scientific Literature using Content Information URL: https://aclweb.org/anthology/W16-5103 [accessed 2019-06-03] [WebCite Cache ID 78r7Ga6MV]
22. Nguyen T, Phung D, Dao B, Venkatesh S, Berk M. Affective and content analysis of online depression communities. IEEE Trans Affective Comput 2014 Jul 1;5(3):217-226. [doi: 10.1109/Taffc.2014.2315623]
23. Monnier J, Laken M, Carter CL. Patient and caregiver interest in internet-based cancer services. Cancer Pract 2002 Nov;10(6):305-310. [doi: 10.1046/j.1523-5394.2002.106005.x]
24. O'Neill B, Ziebland S, Valderas J, Lupiáñez-Villanueva F. User-generated online health content: a survey of internet users in the United Kingdom. J Med Internet Res 2014 Apr 30;16(4):e118 [FREE Full text] [doi: 10.2196/jmir.3187] [Medline: 24784798]
25. Lu Y, Zhang P, Liu J, Li J, Deng S. Health-related hot topic detection in online communities using text clustering. PLoS One 2013 Feb;8(2):e56221 [FREE Full text] [doi: 10.1371/journal.pone.0056221] [Medline: 23457530]
26. Yang F, Lee AJ, Kuo S. Mining health social media with sentiment analysis. J Med Syst 2016 Nov;40(11):236. [doi: 10.1007/s10916-016-0604-4] [Medline: 27663246]
27. Tapi NMD, Bringay S, Lavergne C, Mollevi C, Opitz T. What patients can tell us: topic analysis for social media on breast cancer. JMIR Med Inform 2017 Jul 31;5(3):e23 [FREE Full text] [doi: 10.2196/medinform.7779] [Medline: 28760725]
28. Carelle N, Piotto E, Bellanger A, Germanaud J, Thuillier A, Khayat D. Changing patient perceptions of the side effects of cancer chemotherapy. Cancer 2002 Jul 1;95(1):155-163 [FREE Full text] [doi: 10.1002/cncr.10630] [Medline: 12115329]
29. Sohl SJ, Schnur JB, Montgomery GH. A meta-analysis of the relationship between response expectancies and cancer treatment-related side effects. J Pain Symptom Manage 2009 Nov;38(5):775-784 [FREE Full text] [doi: 10.1016/j.jpainsymman.2009.01.008] [Medline: 19775863]
30. Aslam MS, Naveed S, Ahmed A, Abbas Z, Gull I, Athar MA. Side effects of chemotherapy in cancer patients and evaluation of patients opinion about starvation based differential chemotherapy. J Cancer Ther 2014;05(08):817-822. [doi: 10.4236/jct.2014.58089]
31. Spasić I, Livsey J, Keane JA, Nenadić G. Text mining of cancer-related information: review of current status and future directions. Int J Med Inform 2014 Sep;83(9):605-623 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.06.009] [Medline: 25008281]
32. Reddit. URL: https://www.reddit.com/ [accessed 2019-06-03] [WebCite Cache ID 78r7x3juM]
33. Johnson GJ, Ambrose PJ. Neo-tribes: the power and potential of online communities in health care. Commun ACM 2006 Jan 1;49(1):107-113. [doi: 10.1145/1107458.1107463]
34. RedditBlog. Top Posts of 2013, Stats, and Snoo Year's Resolutions URL: https://redditblog.com/2013/12/31/top-posts-of-2013-stats-and-snoo-years-resolutions/ [accessed 2019-06-03] [WebCite Cache ID 78r86wPbY]
35. Subreddit Stats. /r/cancer stats URL: https://subredditstats.com/r/cancer [accessed 2019-06-03] [WebCite Cache ID 78r8J8oUb]
36. SNOMED. URL: http://www.snomed.org/ [accessed 2019-06-03] [WebCite Cache ID 78rD4S5yd]
37. National Center for Biotechnology Information. Medical Subject Headings URL: https://www.ncbi.nlm.nih.gov/mesh [accessed 2019-06-03] [WebCite Cache ID 78rDzIavh]
38. Gene Ontology (GO) Knowledge Base. URL: http://geneontology.org/ [accessed 2019-06-03] [WebCite Cache ID 78rEGtUWi]
39. GenomeNet. Kyoto Encyclopedia of Genes and Genomes (KEGG) disease database URL: https://www.genome.jp/kegg/disease/ [accessed 2019-06-05] [WebCite Cache ID 78tdhhRi4]
40. DrugBank. URL: https://www.drugbank.ca/ [accessed 2019-06-03] [WebCite Cache ID 78rFjahY6]
41. National Library of Medicine. Unified Medical Language System URL: https://www.nlm.nih.gov/research/umls/ [accessed 2019-06-03] [WebCite Cache ID 78rG0jTjh]

XSL•FO
RenderX

42. TSMM (Text & Social Media Mining) Lab - Yonsei University. PKDE4J URL: http://informatics.yonsei.ac.kr/pkde4j/ [accessed 2019-06-03] [WebCite Cache ID 78rGLBh9h]
43. Stanford CoreNLP. Introduction to pipelines URL: https://stanfordnlp.github.io/CoreNLP/pipelines.html [accessed 2019-06-03] [WebCite Cache ID 78rGV2Gpd]
44. Song M, Kim W, Lee D, Heo G, Kang K. PKDE4J: Entity and relation extraction for public knowledge discovery. J Biomed Inform 2015 Oct:320-332. [Medline: 26277115]
45. Gephi - The Open Graph Viz Platform. URL: https://gephi.org/ [accessed 2019-06-03] [WebCite Cache ID 78rGm1XPj]
46. Wiktionary. Disease URL: https://en.wiktionary.org/wiki/disease [accessed 2019-06-03] [WebCite Cache ID 78rGytBCe]
47. Zappavigna M. Academia.edu. Discourse of Twitter and Social Media: How we use language to create affiliation on the web URL: https://www.academia.edu/18311721/Discourse_of_Twitter_and_Social_Media_How_we_use_language_to_create_affiliation_on_the_web [accessed 2019-06-03] [WebCite Cache ID 78rH6DFwW]
48. Simard S, Savard J. Fear of Cancer Recurrence Inventory: development and initial validation of a multidimensional measure of fear of cancer recurrence. Support Care Cancer 2009 Mar;17(3):241-251. [doi: 10.1007/s00520-008-0444-y] [Medline: 18414902]
49. Groenvold M, Petersen M, Aaronson N, Arraras J, Blazeby J, Bottomley A, et al. EORTC QLQ-C15-PAL: the new standard in the assessment of health-related quality of life in advanced cancer? Palliat Med 2006 Mar;20(2):59-61. [Medline: 16613400]
50. Yanez B, Pearman T, Lis C, Beaumont J, Cella D. The FACT-G7: a rapid version of the functional assessment of cancer therapy-general (FACT-G) for monitoring symptoms and concerns in oncology practice and research. Ann Oncol Off J Eur Soc Med Oncol 2013 Apr;24(4):1073-1078. [Medline: 23136235]
51. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. nitial adult health item banks and first wave testing of the Patient-Reported Outcomes Measurement Information System (PROMIS™) network: 2005–2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
52. Denecke K, Nejdl W. How valuable is medical social media data? Content analysis of the medical web. Inf Sci 2009 May 30;179(12):1870-1880. [doi: 10.1016/j.ins.2009.01.025]
53. Korkontzelos I, Nikfarjam A, Shardlow M, Sarker A, Ananiadou S, Gonzalez GH. Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. J Biomed Inform 2016 Aug;62:148-158 [FREE Full text] [doi: 10.1016/j.jbi.2016.06.007] [Medline: 27363901]

Abbreviations

ALL: acute lymphoblastic leukemia
AML: acute myeloid leukemia
BRCA2: breast cancer type 2
CLL: chronic lymphocytic leukemia
CML: chronic myeloid leukemia
COPD: chronic obstructive pulmonary disease
C-topic: conditional topic
DCIS: ductal carcinoma in situ
FWIW: For What it's Worth
GERD: gastroesophageal reflux disease
LDA: latent Dirichlet allocation
LOL: laughing out loud
MRSA: methicillin-resistant *Staphylococcus aureus*
TSH: thyroid stimulating hormone

Edited by G Eysenbach; submitted 22.11.18; peer-reviewed by A Sarker, A Kulanthaivel, M Waring, E Van Den Broek-Altenburg; comments to author 21.12.18; revised version received 06.05.19; accepted 21.05.19; published 13.06.19.

Please cite as:
Foufi V, Timakum T, Gaudet-Blavignac C, Lovis C, Song M
Mining of Textual Health Information from Reddit: Analysis of Chronic Diseases With Extracted Entities and Their Relations
J Med Internet Res 2019;21(6):e12876
URL: http://www.jmir.org/2019/6/e12876/
doi:10.2196/12876
PMID:31199327

XSL•FO
RenderX

©Vasiliki Foufi, Tatsawan Timakum, Christophe Gaudet-Blavignac, Christian Lovis, Min Song. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Participant Engagement in and Perspectives on a Web-Based Mindfulness Intervention for 9-1-1 Telecommunicators: Multimethod Study

Darragh C Kerr[1], MPH; India J Ornelas[1], MPH, PhD; Michelle M Lilly[2], PhD; Rebecca Calhoun[3], MPH; Hendrika Meischke[1,3], MPH, PhD

[1]Department of Health Services, University of Washington, Seattle, WA, United States

[2]Department of Psychology, Northern Illinois University, DeKalb, IL, United States

[3]Northwest Center for Public Health Practice, University of Washington, Seattle, WA, United States

Corresponding Author:
Darragh C Kerr, MPH
Department of Health Services
University of Washington
Department of Health Services
1959 NE Pacific St
Seattle, WA, Washington
United States
Phone: 1 206 616 4923
Fax: 1 206 543 3964
Email: dckerr@uw.edu

Abstract

Background: Demanding working conditions and secondary exposure to trauma may contribute to a high burden of stress among 9-1-1 telecommunicators, decreasing their ability to work effectively and efficiently. Web-based mindfulness-based interventions (MBIs) can be effective in reducing stress in similar populations. However, low engagement may limit the effectiveness of the intervention.

Objective: The aim of this study was to assess participant engagement in a Web-based MBI designed for 9-1-1 telecommunicators. Specifically, we sought to describe the following: (1) participant characteristics associated with intervention engagement, (2) participant perspectives on engaging with the intervention, and (3) perceived challenges and facilitators to engaging.

Methods: We used qualitative and quantitative data from participant surveys (n=149) that were collected to assess the efficacy of the intervention. We conducted descriptive and bivariate analyses to identify associations between demographic, psychosocial, and workplace characteristics and engagement. We conducted a thematic analysis of qualitative survey responses to describe participant experiences with the MBI.

Results: We found that no individual participant characteristics were associated with the level of engagement (low vs high number of lessons completed). Participant engagement did vary by the call center (P<.001). We identified the following overarching qualitative themes: (1) the participants perceived benefits of mindfulness practice, (2) the participants perceived challenges to engage with mindfulness and the intervention, and (3) intervention components that facilitated engagement. The participants expressed positive beliefs in the perceived benefits of practicing mindfulness, including increased self-efficacy in coping with stressors and increased empathy with callers. The most commonly cited barriers were work-related, particularly not having time to participate in the intervention at work. Facilitators included shorter meditation practices and the availability of multiple formats and types of intervention content.

Conclusions: The findings of this study suggest that efforts to improve intervention engagement should focus on organizational-level factors rather than individual participant characteristics. Future research should explore the effect of mindfulness practice on the efficiency and effectiveness of 9-1-1 telecommunicators at work.

Trial Registration: ClinicalTrials.gov NCT02961621; https://clinicaltrials.gov/ct2/show/NCT02961621

(J Med Internet Res 2019;21(6):e13449) doi:10.2196/13449

XSL•FO
RenderX

KEYWORDS

occupational stress; occupational health; mental health; mindfulness; telecommunications

Introduction

Background

9-1-1 telecommunicators listen to the worst events of our lives, experiencing a slice of a traumatic moment alongside us. Although 9-1-1 telecommunicators are not present at the scene, they may experience secondary trauma. These stressors, coupled with the unpredictability of calls and limited control over the events they are witnessing, may lead to emotional distress [1,2]. Exposure to secondary trauma and work-related stress in 9-1-1 telecommunicators has been associated with posttraumatic stress disorder (PTSD) [3,4], acute stress disorder [5], secondary traumatic stress [6], and occupational burnout [5,6]. The costs of work-related stress extend beyond the individual telecommunicator, including decreased productivity [7], increased absenteeism [7], and increased health care utilization and expenditures [8,9]. With the lives of others depending on the ability of 9-1-1 telecommunicators to work effectively and efficiently, there is a significant need for stress-reduction interventions within this population.

Mindfulness-based interventions (MBIs) have been shown to reduce stress in general populations [10,11]. MBIs encourage participants toward a nonjudgmental acceptance of stressors through the cultivation of present-centered awareness [12-14]. Web-based MBIs, delivered via the internet or a computer portal, have been shown to have comparable effect sizes on stress as in-person formats [15]. Furthermore, Web-based MBIs may be particularly well suited for the 9-1-1 telecommunicator population in which shift work and variable schedules make in-person training difficult. However, low engagement and high dropout rates in Web-based MBIs and other Web-based health promotion interventions can limit intervention effectiveness [16-19]. Several studies and meta-analyses of Web-based interventions have reported a dose-response relationship in which increased engagement is associated with better outcomes [19-21]. An understanding of who engages and how they engage with Web-based MBIs is important to ensure that the interventions are reaching those who may need them most, to maximize engagement and to improve the effectiveness of future interventions. This necessitates a multimethod approach that assesses effective engagement, or "sufficient engagement to achieve the intended outcomes," rather than quantitative measures of engagement alone [22].

Studies have indicated that a paradoxical relationship may exist, in which the psychological symptoms that MBIs intend to ameliorate may make it more difficult for participants first engaging in mindfulness practice to attain present-centered awareness [23-26]. Thus, MBI participants may become discouraged and disengage from the intervention [26]. Stress [24], low trait mindfulness [25], PTSD [25], and perseverative thinking [23,26] have been proposed as factors that may hinder engagement in mindfulness practice. In addition, overcommitment, a perseverative thinking style marked by "excessive striving and a strong need for approval and esteem at work" may make it more difficult for some participants in workplace MBIs to engage [1,27]. Conversely, a commitment to practice and to incorporate mindfulness into daily life, as well as positive beliefs in the benefits of practicing mindfulness, may encourage engagement [13,28,29]. This study sought to explore how such factors influence engagement in MBIs.

Workplace MBIs may have additional barriers or facilitators to participant engagement. Supportive work environments may offer social support, which is a facilitator of engagement [18,30-32]. Factors outside of the control or measure of the intervention, such as staffing policies or the workplace physical environment, may also influence engagement.

Objectives

This study is an analysis of several aspects of participant engagement in a Web-based workplace MBI designed to reduce stress and increase mindfulness in 9-1-1 telecommunicators. In addition to behavioral outcomes (the outcome paper is under review), the randomized controlled trial included process measures related to engagement and qualitative survey data collected throughout the intervention. The aims of this study were to use both the process data and qualitative survey data to explore the following: (1) measurable factors associated with the level of engagement (low vs high) in the intervention, (2) participant perspectives on engaging with the MBI, and (3) the perceived challenges and facilitators to engagement in general.

Methods

Study Design and Recruitment

This study was an analysis of participant engagement in the treatment arm of a randomized controlled trial (ClinicalTrials.gov NCT02961621), testing the effectiveness of a Web-based MBI to reduce stress among 9-1-1 telecommunicators [33]. The outcomes of the efficacy study are currently under review.

Study recruitment was conducted in 2 phases. First, emergency response call centers in the United States and Canada were recruited using industry publications and email announcements. To be eligible for the study, the call centers had to allow their employees to receive emails and use the internet to access the intervention website. The enrolled call centers (n=31) represented rural, urban, and suburban areas in the United States and Canada and responded to 9-1-1 calls for police, fire, or medical emergencies or a combination of call types. In the second phase of recruitment, individual 9-1-1 telecommunicators were recruited from within the enrolled call centers. Recruitment differed at each call center but relied mainly on staff announcements, recruitment emails and flyers, and word-of-mouth. Individual participants were required to be currently employed as a 9-1-1 telecommunicator (ie, call-receivers, dispatchers, or both). A total of 323 telecommunicators (n=325 were assessed for eligibility; n=2 declined to participate) were enrolled in the study [33].

XSL•FO
RenderX

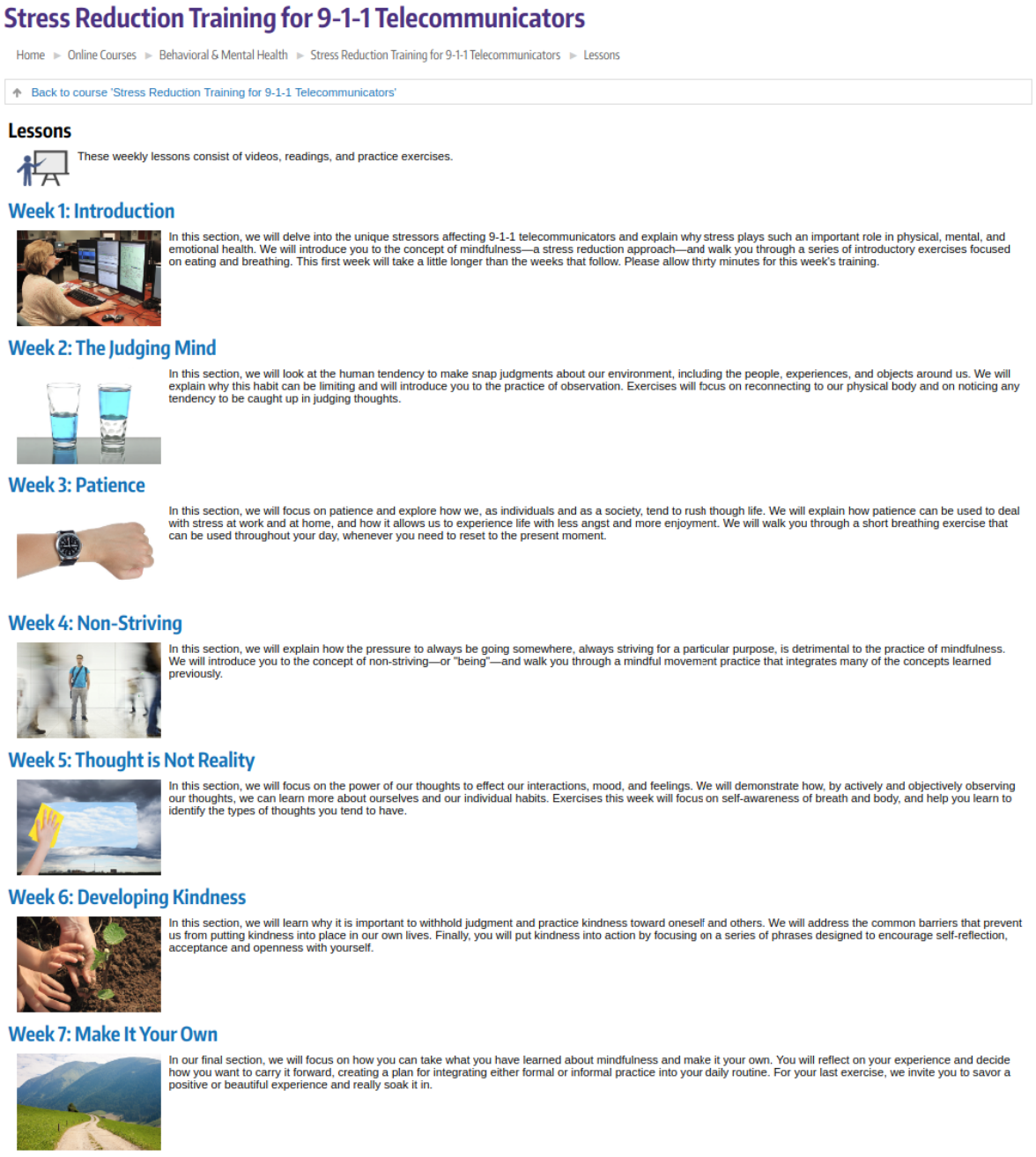
The study was approved by the Institutional Review Board of the University of Washington. After obtaining electronic informed consents, the participants provided demographic and employment-related (ie, length of employment) information and completed a Web-based baseline survey [33].

Web-Based Mindfulness Intervention

The Web-based mindfulness intervention was modeled after Mindfulness-Based Stress Reduction, which has been shown to be effective for a variety of physical and mental conditions [10,11,33-35]. Clinicians and investigators trained in mindfulness developed the intervention to meet the specific needs of 9-1-1 telecommunicators. After consultation with stakeholders at the enrolled call centers, intervention developers adapted the intervention content from the traditional in-person format to an abbreviated Web-based format to address logistical concerns.

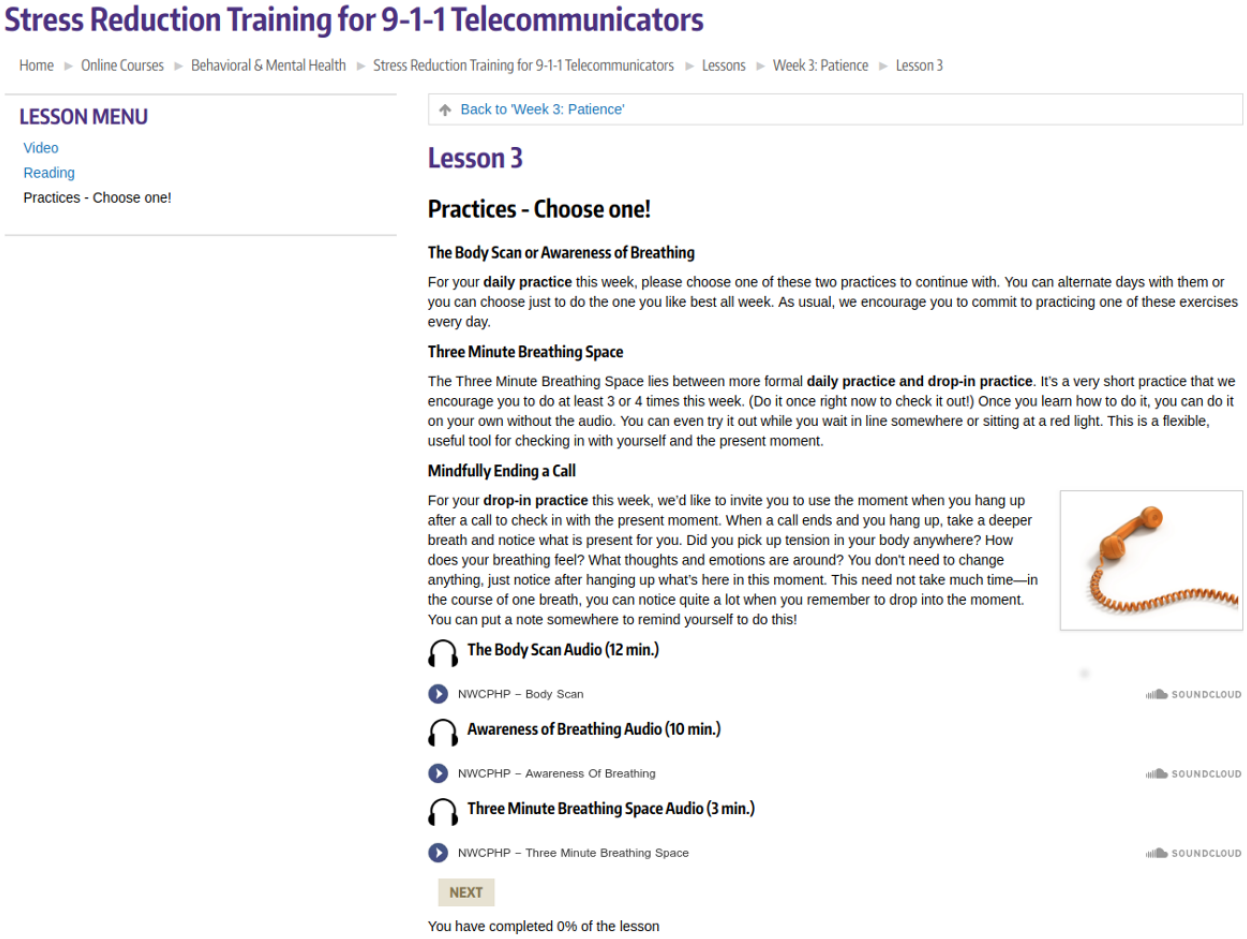
The intervention's 7 Web-based lessons were hosted on the learning management system of the Northwest Center for Public Health Practice at the University of Washington (Figures 1 and 2). Each lesson started with a short video that introduced that week's theme and was followed by a short reading. The next section of the lessons consisted of 1 longer (10 to 14 min) *daily practice* with guided audio that introduced formalized meditation skills, such as body scan and loving-kindness, and 1 to 2 brief *drop-in* mindfulness practices focused on incorporating mindfulness activities into daily life. Some of these practices, such as *body awareness at your desk* and *mindfully ending a call*, were tailored specifically for the emergency response call center environment. Each lesson also included a weekly check-in survey and an optional moderated discussion board. The estimated time to complete each lesson was between 20 and 30 minutes [33].

Figure 1. Overview of weekly lessons.



XSL•FO
RenderX

Figure 2. Suggested daily and drop-in practices for 1 lesson.



Intervention Procedures

After completion of a baseline survey, the participants randomized to the intervention (N=161) were contacted twice weekly throughout the intervention period. One email contained a link to the weekly training lesson, whereas the second email provided suggestions for incorporating mindfulness skills into daily life.

Call center managers were highly encouraged to provide the study participants with a designated time during work to complete the intervention. Participants were asked to complete 1 lesson per week over a 7-week period and were encouraged to complete the lessons on a designated weekday as their work schedules allowed. However, lessons from previous weeks could be accessed throughout the intervention period. Participants were instructed to do the *daily practice* with guided audio for approximately 10 minutes for at least 6 out of 7 days a week and were encouraged to do the *drop-in* practices as often as they were able to.

At the beginning of each weekly lesson, intervention participants were asked to complete Web-based weekly check-in surveys with a mix of close- and open-ended questions. These surveys assessed how often the participants practiced formal mindfulness during the previous week, if and how they incorporated mindfulness into their daily lives, and any perceived effects. There were 6 weekly check-in surveys in total. A final training evaluation after the last weekly lesson (week 7) assessed participant satisfaction, perceived effects, and overall experience. Textbox 1 lists the open-ended questions.

Web-based follow-up surveys identical to the initial baseline survey were completed at 2 weeks and 4 months after the end of the intervention. All the participants received a certificate of completion, and the Washington State Criminal Justice Training Commission: Telecommunicator Program recognized the training as continuing education for renewal of telecommunicator certification [33].

Measures

Main Outcome

Engagement

Engagement in the intervention was measured as the number of lessons completed, assessed via access logs for viewing lesson videos. Owing to the small number of participants per lesson completed, we categorized participants into 2 groups: those who completed 0 to 4 lessons (low engagement) and those who completed 5 to 7 lessons (high engagement).

Demographic

Participants were asked to self-report their age, gender, race, ethnicity, marital status, highest level of education, years of experience as a 9-1-1 telecommunicator, and whether, currently, they had children under the age of 18 years.

XSL•FO
RenderX

Textbox 1. Open-ended survey questions used in qualitative analysis.

Weekly check-in survey

1. Can you give 1 or 2 examples of how you incorporated mindfulness into your daily life this week?
2. Please share anything you have noticed this week about the effects of your mindfulness practice.

Final training evaluation

1. What did you like about this training?
2. What effect has this training had on your stress level, if any?
3. What did you dislike or what would you change about the training?
4. Is there anything else you would like to share with us about your experience with this training?

Psychosocial

The Calgary Symptoms of Stress Inventory

Stress was measured using the 56-item Calgary Symptoms of Stress Inventory, which assesses the frequency of experiencing subjective symptoms of stress with 8 subscales. Participants rate the frequency of experiencing the stress-related symptoms on a 5-point Likert scale from *never* to *frequently* during the past week [36]. Previous research has established the reliability of the instrument in a sample of 9-1-1 telecommunicators, and internal consistency of the scale in the main outcome study was strong (alpha=.95) [1, personal communication by H Meischke, December 5, 2018].

Mindful Attention Awareness Scale

Mindfulness was measured using the 15-item Mindful Attention Awareness Scale (MAAS), which assesses mindfulness as a unidimensional construct of attentional awareness in the present moment. Participants are asked how frequently or infrequently they have each stated the experience of mindlessness, conceptualized as the inverse of mindfulness, on a 6-point Likert scale from *almost always* to *almost never* [37]. The MAAS has been used in previous research with 9-1-1 telecommunicators, and the internal consistency of the scale in the main outcome study was strong (alpha=.91) [1, personal communication by H Meischke, December 5, 2018].

Work-Related

Swedish Demand-Control-Support Questionnaire Social Support Subscale

The 6-item social support subscale of the Demand-Control-Support Questionnaire (DCSQ) was used to assess overall workplace atmosphere and social support from coworkers and supervisors. Participants are asked to report their agreement with statements on a 4-point Likert scale ranging from *strongly disagree* to *strongly agree* [38]. Previous research has established good internal consistency, construct validity, and criterion validity for the English version of the total DCSQ scale (alpha=.83) and good internal consistency for the social support subscale (alpha=.84) [39].

Social Support Visual Analog Scale

A visual analog scale (VAS) was also used to measure the overall level of satisfaction with social support in the workplace. The scale ranges from 0 to 100 with 0 representing *completely dissatisfied with social support at work* and 100 representing *completely satisfied with social support at work*. The scale has been employed in previous research in the emergency responder population [40].

Network Conflict Visual Analog Scale

A VAS was used to measure the perceived degree of conflict in the individual participant's workplace social network. The scale ranges from 0 to 100 with 0 representing *little or no conflict with coworkers* and 100 representing *frequent and intense conflict with coworkers*. The scale has been employed in previous research in the emergency responder population [40].

Effort-Reward-Imbalance Overcommitment Subscale

The 6-item overcommitment subscale of the effort-reward-imbalance (ERI) is used to measure an individual's tendency to engage in a coping pattern of excessive commitment and the high need for approval at work [41]. Responses are indicated on a 4-point Likert scale with higher scores indicating greater overcommitment. The subscale has sound psychometric properties, including satisfactory confirmatory factor analysis, internal consistency, and reliability [42]. The ERI has been employed in previous research with 9-1-1 telecommunicators, which found good internal consistency for the overcommitment subscale (alpha=.85) [1].

Data Analysis

We used a complementary integration of qualitative and quantitative methods to suit each inquiry in this analysis [43]. We employed 2 research paradigms, quantitative positivist and qualitative interpretivist, and have provided a convergent interpretation of the results in the Discussion section [44].

Quantitative Data Analysis

Quantitative methods were used to describe demographic, psychosocial, and workplace characteristics associated with the intervention engagement. Statistical analyses were conducted using the R statistical package version 3.4.3 [45]. Bivariate associations between the independent variables (ie, demographic, psychosocial, and workplace characteristics) and participant engagement (ie, low vs high engagement) were tested using chi-square or Fisher exact tests for categorical variables and unpaired *t* tests for continuous variables.

XSL•FO
RenderX

Qualitative Data Analysis

Qualitative methods were used to describe the participants' perspectives of the intervention and to identify perceived barriers and facilitators. A thematic analysis was conducted using Web-based responses to 6 open-ended questions from 6 weekly check-in surveys and a final training evaluation (Textbox 1) [46]. In total, 822 open-ended responses from 109 unique participants at 26 call centers were analyzed. The analysis was conducted without the use of the high-low engagement characterization used in the quantitative analysis.

The first author read the entirety of the data to become familiar with the content and then independently developed an initial coding schema through a hybrid deductive-inductive approach [47]. Deductive coding was used to create an initial codebook that was derived from Banerjee et al's 5-facet framework for psychological engagement (ie, motivation, intention, commitment, belief, and a therapeutic relationship between the individual, teacher, and group) [23]. On the basis of the content of the survey questions, only 2 out of the 5 facets, commitment and belief, were included in the initial codebook. Meanwhile, inductive coding allowed for the inclusion of themes outside the bounds of the theoretical framework. After creation of the pilot codebook based on the deductive approach, additional broad code categories and subcodes were added based on an initial read-through. A second coder unaffiliated with the study was trained on the coding schema; then the 2 coders independently piloted the schema to establish an initial agreement on how to apply the codes. The 2 team members met iteratively to verbally negotiate any discrepancies during coding and to ensure credibility of the analysis. The first coder identified major themes, and the second coder reviewed the themes to ensure credibility. Sample quotations were identified to illustrate each theme.

Results

Participant Characteristics and Engagement

Of the 161 participants enrolled in the intervention's treatment arm, 12 (7.5%) were excluded from this analysis owing to loss of employment or not completing the baseline survey. Table 1 portrays the sample characteristics of the participants in this analysis (n=149). Nearly half of the participants (71/149, 47.7%) completed all 7 intervention lessons whereas 21.5% (32/149) did not complete a single lesson. The mean number of lessons completed was 4.7 (SD 2.8). The mean number of days the participants reported practicing formal mindfulness with guided audio (3 [SD 0.15]) did not differ by engagement group (low vs high) and remained consistent through the intervention.

Of all the participants, 35.6% (53/149) were characterized as low engagement (0 to 4 lessons complete) and 64.4% (96/149) as high engagement (5 to 7 lessons complete). The comparison between low and high engagement showed that demographic, psychosocial, or work-related characteristics were not significantly related to engagement (data not shown). Only the call center of employment was statistically significantly associated with engagement (P<.001).

Participant Perspectives on the Mindfulness-Based Intervention

A total of 3 overarching themes were identified: (1) the participants perceived benefits of mindfulness practice, (2) the participants perceived challenges to engaging with mindfulness and the intervention, and (3) intervention components that facilitated engagement. Textbox 2 provides an outline of the themes and subthemes.

The Participants Perceived Benefits of Mindfulness Practice

This theme describes the perceived benefits of engaging in mindfulness practice. Overall, the participants indicated that with continued practice and incorporation of mindfulness into their daily life, practicing mindfulness became easier and the benefits of practice became more perceptible. Consequently, many of the participants indicated that they were more able and willing to continue:

> *...[mindfulness is] becoming easier to incorporate into my everyday. The more I practice it, the more I want to do it as I am noticing it's helpful to relax.*

Mindfulness Helps 9-1-1 Telecommunicators Cope With Stress

The participants frequently reported uncertainty of a direct effect on their stress level; instead, many noted that practicing mindfulness positively changed how they coped with stressors. One participant expressed:

> *I am not suddenly Zen, but I feel I can identify stress quicker and shake it off easier.*

Many of the participants reported that although their stress was still present, they felt more aware, accepting, and in control of how their stress affected their mental and physical well-being:

> *It's almost like the stress is still hovering there, but I can choose to not think about the stress.*

These responses indicate a positive belief in not only the benefits of mindfulness practice but also in one's ability to apply the concepts of mindfulness to reduce stress.

Mindfulness Helps 9-1-1 Telecommunicators Communicate With Callers and Focus at Work

Many of the participants reported that mindfulness enabled them to have more empathy with callers:

> *I have been trying to change how I think of our callers that frustrate me. Instead of thinking that they are all stupid, I am trying to think they are being silly, or to be more empathetic.*

XSL•FO
RenderX

Table 1. Characteristics of participants.

Variables	Participants (N=149)[a]
Demographic characteristics, n (%)	
Age (years)	
Below 26	11 (7.4)
18 to 35	50 (33.6)
36 to 45	51 (34.2)
46 to 55	28 (18.8)
56 to 64	9 (6.0)
Gender	
Female	126 (84.6)
Male	23 (15.4)
Race	
American Indian or Alaska Native	7 (4.3)
Asian	1 (0.6)
Black	4 (2.4)
Multiracial	7 (4.3)
Native Hawaiian or Pacific Islander	0 (0.0)
Other	5 (3.0)
White	141 (86.0)
Binary race	
Nonwhite	8 (5.4)
White	141 (94.6)
Ethnicity	
Hispanic	4 (2.8)
Non-Hispanic	141 (97.2)
Experience as a 9-1-1 telecommunicator (years)	
Less than 2	20 (13.5)
2 to 10	68 (45.9)
11 to 20	42 (28.4)
Above 20	18 (12.2)
Married or living with a partner	
Yes	97 (65.5)
No	51 (34.5)
Children under the age of 18 years	
Yes	67 (45.0)
No	82 (55.0)
Highest education	
High school or General Education Diploma	12 (8.1)
Some college	63 (42.3)
Associates	14 (9.4)
Bachelors	51 (34.2)
Postgraduate study or degree	9 (6.0)
Psychosocial characteristics, mean (SD)	

XSL•FO
RenderX

Variables	Participants (N=149)[a]
Stress (Calgary Symptoms of Stress Inventory)	56.0 (27.5)
Mindfulness (Mindful Attention Awareness Scale)	4.1 (0.8)
Work-related characteristics, mean (SD)	
Social support (VAS[b])	67.8 (24.0)
Social support (Demand-Control-Support Questionnaire subscale)	17.6 (2.9)
Network conflict (VAS)	33.3 (25.0)
Overcommitment (effort-reward-imbalance subscale)	13.4 (3.9)

[a]Some frequencies do not sum up to 149 owing to missing responses.
[b]VAS: visual analog scale.

Textbox 2. Overview of qualitative findings: themes and subthemes.

Theme 1: The participants perceived benefits of mindfulness practice

- Mindfulness helps 9-1-1 telecommunicators cope with stress
- Mindfulness helps 9-1-1 telecommunicators communicate with callers and focus at work
- The participants perceived broad benefits of mindfulness in their everyday life

Theme 2: The participants perceived challenges in engaging with mindfulness and the intervention

- Stress and mindlessness make it difficult to attain present-centered awareness
- Aspects of the call center environment make it difficult to engage

Theme 3: Intervention components facilitated engagement

- Shorter mindfulness practices were easier to engage with and use at work
- Variety in intervention content allowed participants to choose what worked best for them

Some reported that mindfulness improved their overall communication on the job:

> *When you approach someone from a place of sincere kindness, it is harder for them to be rude, angry or hostile. I talked to many people on the phones this week very escalated and upset with the police response time, and by using this approach and using my breath to remain aware and present, and non-reactive, I was able to establish better rapport with these callers.*

In addition, some of the participants reported that practicing mindfulness improved their focus at work, especially when multitasking. As one participant stated:

> *In our career we have to multitask all the time...it can be easy to overlook something. I tried to put more of my focus into the audible while still multitasking but noticed I was catching things clearer.*

The Participants Perceived Broad Benefits of Mindfulness in Their Everyday Life

Participants described how being mindful improved their overall quality of life. Feeling calm was one of the most commonly reported perceived benefits of practicing mindfulness. Many participants attributed this sense of calm to an increased acceptance of external situations. As one participant described it:

> *I feel calmer. I try not to fret about things that I cannot change.*

Many participants also noted that practicing mindfulness made them feel relaxed, energized, and happier. For some of the participants, bringing present-centered awareness into their daily routines made them feel less rushed, thus increasing their appreciation of everyday life. One participant expressed that:

> *Rather than rushing from task to task I became more present and slowed down, and enjoyed the process of even mundane tasks like folding laundry or putting away groceries.*

Some participants noted that practicing mindfulness expanded their self-awareness and increased their sense of control over thoughts:

> *I think it has made me more aware of how I'm feeling and how my feelings are affecting me both physically and mentally.*
>
> *I feel more in control of what I'm thinking, acknowledging the randomness of some thoughts.*

Participants frequently reported the perceived consequences of practicing mindfulness on their interpersonal relationships as well. Many participants described becoming more patient and less emotionally reactive, for example:

XSL•FO
RenderX

> *I was able to have a difficult conversation with a family member (parent) much more calmly and less emotionally than in the past. I feel it was different this time because of the mindfulness practices.*
>
> *I'm finding myself to be way less reactive, I'm more serene and calm, and less sensitive. I feel I'm rolling with the punches and not letting emotional or angry callers get a rise out of me.*

In addition, some participants noted that practicing mindfulness improved the quality of their relationships with loved ones:

> *My kids are excited to tell me about their day because I'm taking more time to focus (that's hard for me sometimes) and really hear them. They love seeing me excited about "their" day.*

Finally, participants frequently reported using mindfulness to cope with physical stressors. Most commonly, participants described practicing mindfulness to fall asleep. Some reported that this helped them not only to fall asleep more quickly but also improved the quality and quantity of sleep.

The Participants Perceived Challenges to Engage With Mindfulness and the Intervention

This theme describes the perceived individual, workplace, and intervention-related challenges to engage with mindfulness and the intervention. Although these challenges may not have led the participants to stop using the intervention, they may have made it more difficult for the participants to engage.

Stress and Mindlessness Make it Difficult to Attain Present-Centered Awareness

Although many participants described using mindfulness to cope with stressors, some participants described being in a stressful state as a challenge to practicing mindfulness:

> *[Practicing mindfulness] was difficult this week as my stress and anxiety level was much higher this week...I tried to be mindful as a way of reducing that, but it was difficult.*

Being in a stressful state made it more difficult for some participants to attain present-centered awareness. As one participant described it:

> *I still find during stressful events at work I lose this feeling of awareness and hours can go by where I haven't focused on my breathing or awareness once.*

Some participants described difficulty in disengaging from habitual mindlessness. For example:

> *I still find that I am going through the motions of my daily activities without thinking about them. Reflection is an afterthought.*

However, this may represent a growing awareness of mindfulness and present-centered thinking rather than a challenge to it.

Aspects of the Call Center Work Environment Make it Difficult to Engage

Some participants described physical and psychological aspects of the work environment that made it difficult to engage with the intervention, especially with the guided audio content. Frequent interruptions, noise levels, shared workspaces, and busy working conditions were often cited:

> *I did not like the mindfulness audio. I was not able to use it while at work due to how busy it was, and it was much easier to practice it without the audio on my own time.*

Some of the participants posited that aspects of the work environment may have limited the effectiveness of the training:

> *It's a difficult training to get done in our center due to shared computers and busy shifts. It's not possible to do the segments in our center without interruption and I think this causes it not to be as effective.*

Participants from nearly 50% of call centers in the intervention reported not having a designated time-off to complete the intervention despite the intervention protocol. This was a major barrier for completing the intervention content, and some participants believed the intervention increased rather than reduced their stress. For example:

> *I am finding it is almost causing more stress trying to find the time to get practice in and to do the weekly lessons. We do not have the staffing to permit us time off the floor to complete training, so we must do it while on duty on the floor.*

Some participants reported completing the intervention lessons while actively working or while at home. As one participant reported:

> *It was difficult to listen to the longer listening exercises at our desks while still answering calls and radio traffic. At times, there were too many interruptions that I would get frustrated and just do it at home.*

A few participants expressed a belief that mindfulness was incongruous with their work as a 9-1-1 telecommunicator:

> *When it is busy on the dispatch floor, there really isn't time to stop and do anything for yourself for the calls that stress you out. Dispatching isn't that kind of job.*
>
> *I think that with our job, we do so much multitasking that it becomes hard to focus on just one thing.*

Intervention Components Facilitated Engagement

This theme describes aspects of the intervention that the participants indicated made it easier for them to engage.

Shorter Mindfulness Practices Were Easier to Engage With and Use at Work

Many participants described shorter-length practices as easier to engage with than longer practices. This was often attributed to difficulty in maintaining present-centered awareness during longer practices:

> *I found it hard to concentrate for the longer exercises, like 12-14 minutes. I know [Name of Instructor] said it was okay to wander, that is part of the process, but I wandered a lot - and kept checking the timer.*

XSL•FO
RenderX

Shorter practices were described as "an attainable goal" in comparison with longer practices. Many of the participants reported difficulty in finding the time to do the longer practices at work owing to frequent interruptions. Similarly, some of the participants felt that they would be more able to use shorter practices at work:

> *We are used to short breaks, always being in a rush... short 5-minute meditations are much more realistic and practical.*

Variety in Intervention Content Allowed Participants to Choose What Worked Best for Them

Some participants noted that they enjoyed having a variety of content formats (ie, audio, video, and written) and types of practice (eg, loving-kindness and body scan) to try each week. The availability of different content allowed participants to find what worked best for them. As one participant shared:

> *I liked that there was a variety of practices to try. Different things work for different people and that was taken into account.*

Discussion

Principal Findings

The goal of this study was to explore the engagement of participants in a Web-based workplace MBI for 9-1-1 telecommunicators with the aims of (1) investigating demographic, psychosocial, and workplace characteristics associated with level of engagement (low vs high) in the intervention, (2) exploring participants' perspectives on engaging in the intervention, and (3) identifying perceived challenges and facilitators to engaging in the intervention. Results of the quantitative analysis showed that individual-level demographic, psychosocial, or workplace characteristics examined in this study were not associated with engagement. Call center location was the only factor associated with engagement, indicating that workplace-level characteristics rather than individual-level characteristics may be more relevant to intervention engagement. The qualitative results supported this finding; difficulty engaging with the intervention at work was commonly identified as a barrier for participants. In nearly 50% of call centers, at least one participant described not having a designated time to complete the intervention. When developing future MBIs for high-stress work environments, researchers should work closely with the workplace managerial staff to identify potential barriers to employee engagement and develop practical solutions to overcome these barriers. The availability of different formats of intervention materials (eg, written, video, and audio) and shorter practices may be simple mechanisms to facilitate engagement in workplaces that are not able to provide employees time-off to engage with the intervention content.

Although we found no significant association between baseline stress and engagement in the intervention, our qualitative findings suggest that being in a stressful state of mind may make it more difficult to attain present-centered awareness. Moreover, the perception of being time-poor may discourage engagement, especially for workplace interventions that may be perceived as detracting from work time [16]. Although more research is needed on the effectiveness of abbreviated meditation practices, it is nevertheless important to consider that an intervention that is used is more effective than one that is not. Furthermore, more attention should be placed on attaining sufficient engagement to elicit behavior change rather than greater quantities of engagement alone [22]. Brief interventions and short practices may act as stepping-stones to more rigorous practice.

Despite the barriers to engage with the intervention, participants exhibited a high degree of commitment to incorporate mindfulness into their daily lives and expressed positive beliefs in the perceived benefits of practicing mindfulness. Previous findings suggest a positive feedback loop for mindfulness practice in which practice prompts greater perceived benefits, which in turn increases motivation to practice [24]. Although mindfulness is inherently not outcome-focused, future MBIs could highlight the positive benefits of mindfulness practice to encourage engagement.

This intervention's main outcome study found significant improvements in stress scores postintervention [personal communication by H Meischke, December 5, 2018]. Our qualitative findings help elucidate how participants perceived stress reduction during the intervention. Specifically, participants reported that although their stress was still present, they felt more aware, accepting, and in control of how their stress affected their mental and physical well-being. Mindfulness practice enabled participants to reclassify stress as a changeable rather than a static state of mind while also increasing their efficacy to enact change. These findings suggest that workplace MBIs can provide participants with tools to cope with stressors in a more effective way.

The benefits of mindfulness practice may extend beyond the individual practicing; participants reported that mindfulness increased their focus at work and their empathy with callers thereby improving their overall communications. Future research is needed to determine the effect of mindfulness practice on the efficiency and effectiveness of 9-1-1 telecommunicators. However, this research suggests that mindfulness may be a unique tool for improving communications, emergency or otherwise.

Strengths and Limitations

A key strength of this study was the use of both qualitative and quantitative methods, which allowed for a richer exploration of intervention engagement than possible with either method alone. Nonetheless, several limitations warrant discussion. Although our study showed a significant difference in participant engagement between call centers and the participants perceived aspects of the workplace as a challenge to engage, we were unable to determine a causal mechanism between organizational-level factors and participant engagement. Additional research using a randomized trial design would inform organizational-level comparisons of participant engagement. Second, we did not have a reliable measure of the intensity of use or attrition. Future studies should include measures that can validly and reliably assess the intensity of use and other process outcomes. Third, the use of Web-based survey data limited the breadth and depth of the thematic analysis. Future qualitative studies using semistructured

XSL•FO
RenderX

participant interviews could identify novel themes or add nuance to this study's findings. Finally, the demographic homogeneity of this study population may limit the external validity of our results.

Conclusions

This study explored participant engagement in a Web-based workplace MBI among 9-1-1 telecommunicators. Our findings suggest that organizational-level factors may be more pertinent to engagement than individual participant characteristics. Some of the qualitative feedback presented in this study may inform future MBIs in similar high-stress work environments, particularly the need to work closely with managerial staff at intervention sites to mitigate potential barriers to employee engagement. The availability of multiple formats of intervention materials (eg, written, video, and audio) and shorter meditation practices may be simple, cost-effective mechanisms to facilitate engagement. A Web-based workplace MBI can provide emergency responders with tools to cope with stressors, improve general well-being, and build empathy with callers. Future research should explore the effect of mindfulness practice on the efficiency and effectiveness of 9-1-1 telecommunicators at work.

Acknowledgments

The authors would like to thank Cailin Henley for assisting with the qualitative analysis and Ian Painter for guidance on the quantitative analysis. This study was supported by grant number ROH010536 funded by the National Institute for Occupational Safety and Health and the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services. The authors acknowledge the people—past, present, and future—of the Dkhw'Duw'Absh, the Duwamish Tribe, the Muckleshoot Tribe, and other tribes on whose traditional lands they study and work.

Conflicts of Interest

None declared.

References

1. Meischke H, Painter I, Lilly M, Beaton R, Revere D, Calhoun B, et al. An exploration of sources, symptoms and buffers of occupational stress in 9-1-1 emergency call centers. Ann Emerg Dispatch Response 2015;3(2):28-35 [FREE Full text]
2. Pierce H, Lilly MM. Duty-related trauma exposure in 911 telecommunicators: considering the risk for posttraumatic stress. J Trauma Stress 2012 Apr;25(2):211-215. [doi: 10.1002/jts.21687] [Medline: 22467384]
3. Lilly MM, Allen CE. Psychological inflexibility and psychopathology in 9-1-1 telecommunicators. J Trauma Stress 2015 Jun;28(3):262-266. [doi: 10.1002/jts.22004] [Medline: 25964163]
4. Regehr C, LeBlanc VR, Barath I, Balch J, Birze A. Predictors of physiological stress and psychological distress in police communicators. Police Pract Res 2013 Dec;14(6):451-463. [doi: 10.1080/15614263.2012.736718]
5. Trachik B, Marks M, Bowers C, Scott G, Olola C, Gardett I. Is dispatching to a traffic accident as stressful as being in one? Acute stress disorder, secondary traumatic stress, and occupational burnout in 911 emergency dispatchers. Ann Emerg Dispatch Response 2015;3(1):27-38 [FREE Full text]
6. Troxell RM. PQDT Open. 2008. Indirect exposure to the trauma of others: The experiences of 9-1-1 telecommunicators URL: https://pqdtopen.proquest.com/doc/304351154.html?FMT=ABS [accessed 2019-06-05] [WebCite Cache ID 78tX0Mbp0]
7. O'Keefe LC, Brown KC, Christian BJ. Policy perspectives on occupational stress. Workplace Health Saf 2014 Aug 20;62(10):432-438. [doi: 10.3928/21650799-20140813-02] [Medline: 25139784]
8. Azagba S, Sharaf M. Psychosocial working conditions and the utilization of health care services. BMC Public Health 2011 Aug 11;11:642 [FREE Full text] [doi: 10.1186/1471-2458-11-642] [Medline: 21834974]
9. Goetzel RZ, Anderson DR, Whitmer RW, Ozminkowski RJ, Dunn RL, Wasserman J, Health Enhancement Research Organization (HERO) Research Committee. The relationship between modifiable health risks and health care expenditures. An analysis of the multi-employer HERO health risk and cost database. J Occup Environ Med 1998 Oct;40(10):843-854. [Medline: 9800168]
10. Khoury B, Sharma M, Rush S, Fournier C. Mindfulness-based stress reduction for healthy individuals: a meta-analysis. J Psychosom Res Jun 2015;78(6):519-528. [Medline: 25818837]
11. Chiesa A, Serretti A. Mindfulness-based stress reduction for stress management in healthy people: a review and meta-analysis. J Altern Complement Med 2009 May;15(5):593-600. [doi: 10.1089/acm.2008.0495] [Medline: 19432513]
12. Allexandre D, Bernstein AM, Walker E, Hunter J, Roizen MF, Morledge TJ. A web-based mindfulness stress management program in a corporate call center: a randomized clinical trial to evaluate the added benefit of onsite group support. J Occup Environ Med 2016 Mar;58(3):254-264 [FREE Full text] [doi: 10.1097/JOM.0000000000000680] [Medline: 26949875]
13. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clin Psychol Sci Prac 2003;10(2):144-156. [doi: 10.1093/clipsy/bpg016]

XSL•FO
RenderX

14. Smith BW, Ortiz JA, Steffen LE, Tooley EM, Wiggins KT, Yeater EA, et al. Mindfulness is associated with fewer PTSD symptoms, depressive symptoms, physical symptoms, and alcohol problems in urban firefighters. J Consult Clin Psychol 2011 Oct;79(5):613-617. [doi: 10.1037/a0025189] [Medline: 21875175]
15. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
16. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. J Med Internet Res 2017 Jul 26;19(7):e271 [FREE Full text] [doi: 10.2196/jmir.7583] [Medline: 28747293]
17. 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]
18. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 steps Australia. J Med Internet Res 2015;17(7):e176 [FREE Full text] [doi: 10.2196/jmir.4339] [Medline: 26180040]
19. Strecher VJ, McClure J, Alexander G, Chakraborty B, Nair V, Konkel J, et al. The role of engagement in a tailored web-based smoking cessation program: randomized controlled trial. J Med Internet Res 2008;10(5):e36 [FREE Full text] [doi: 10.2196/jmir.1002] [Medline: 18984557]
20. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
21. Glasgow RE, Christiansen SM, Kurz D, King DK, Woolley T, Faber AJ, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. J Med Internet Res 2011;13(1):e9 [FREE Full text] [doi: 10.2196/jmir.1391] [Medline: 21371992]
22. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
23. Banerjee M, Cavanagh K, Strauss C. Barriers to mindfulness: A path analytic model exploring the role of rumination and worry in predicting psychological and physical engagement in an online mindfulness-based intervention. Mindfulness (N Y) 2018;9(3):980-992 [FREE Full text] [doi: 10.1007/s12671-017-0837-4] [Medline: 29875884]
24. Langdon S, Jones FW, Hutton J, Holttum S. A grounded-theory study of mindfulness practice following mindfulness-based cognitive therapy. Mindfulness 2011 Sep 8;2(4):270-281. [doi: 10.1007/s12671-011-0070-5]
25. Liu B, Rice V. Characteristics of United States military personnel and veterans who complete mindfulness training. Cham, Switzerland: Springer International Publishing AG; 2017 Presented at: AHFE 2017 International Conferences on Human Factors and Ergonomics in Healthcare and Medical Devices; July 17-21, 2017; Los Angeles p. 22-32. [doi: 10.1007/978-3-319-60483-1_3]
26. Crane C, Williams JM. Factors associated with attrition from mindfulness-based cognitive therapy in patients with a history of suicidal depression. Mindfulness (N Y) 2010 Mar 11;1(1):10-20 [FREE Full text] [doi: 10.1007/s12671-010-0003-8] [Medline: 21125023]
27. de Jonge JJ, Bosma H, Peter R, Siegrist J. Job strain, effort-reward imbalance and employee well-being: a large-scale cross-sectional study. Soc Sci Med 2000 May;50(9):1317-1327. [Medline: 10728851]
28. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. J Behav Med 2008 Feb;31(1):23-33. [doi: 10.1007/s10865-007-9130-7] [Medline: 17899351]
29. Banerjee M, Cavanagh K, Strauss C. A qualitative study with healthcare staff exploring the facilitators and barriers to engaging in a self-help mindfulness-based intervention. Mindfulness (N Y) 2017;8(6):1653-1664 [FREE Full text] [doi: 10.1007/s12671-017-0740-z] [Medline: 29201248]
30. Balk-Møller NC, Larsen TM, Holm L. Experiences from a web- and app-based workplace health promotion intervention among employees in the social and health care sector based on use-data and qualitative interviews. J Med Internet Res 2017 Oct 19;19(10):e350 [FREE Full text] [doi: 10.2196/jmir.7278] [Medline: 29051133]
31. Poirier J, Cobb NK. Social influence as a driver of engagement in a web-based health intervention. J Med Internet Res 2012;14(1):e36 [FREE Full text] [doi: 10.2196/jmir.1957] [Medline: 22356829]
32. Robroek SJ, van Lenthe FJ, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. Int J Behav Nutr Phys Act 2009;6:26 [FREE Full text] [doi: 10.1186/1479-5868-6-26] [Medline: 19457246]
33. Meischke H, Lilly M, Beaton R, Calhoun R, Tu A, Stangenes S, et al. Protocol: a multi-level intervention program to reduce stress in 9-1-1 telecommunicators. BMC Public Health 2018 Dec 2;18(1):570 [FREE Full text] [doi: 10.1186/s12889-018-5471-0] [Medline: 29716576]
34. Baer R. Mindfulness training as a clinical intervention: a conceptual and empirical review. Clin Psychol Sci Prac 2003;10:125-143. [doi: 10.1093/clipsy.bpg015]

XSL•FO
RenderX

35. Kabat-Zinn J. Full Catastrophe Living (Revised Edition): Using The Wisdom Of Your Body And Mind To Face Stress, Pain, And Illness. New York: Bantam; 2013.
36. Carlson LE, Thomas BC. Development of the Calgary Symptoms of Stress Inventory (C-SOSI). Int J Behav Med 2007;14(4):249-256. [doi: 10.1080/10705500701638484] [Medline: 18001241]
37. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003 Apr;84(4):822-848. [Medline: 12703651]
38. Sanne B, Torp S, Mykletun A, Dahl A. The Swedish Demand-Control-Support Questionnaire (DCSQ): factor structure, item analyses, and internal consistency in a large population. Scand J Public Health 2005;33(3):166-174. [Medline: 16040456]
39. Mauss D, Herr R, Theorell T, Angerer P, Jian L. Validating the Demand Control Support Questionnaire among white-collar employees in Switzerland and the United States. J Occup Med Toxicol 2018;13(7):eCollection. [Medline: 29449870]
40. Beaton R, Murphy S, Pike K, Corneil W. Social support and network conflict in firefighters and paramedics. West J Nurs Res 1997;19(3):297-313. [Medline: 9170989]
41. Siegrist J, Starke D, Chandola T, Godin I, Marmot M, Niedhammer I, et al. The measurement of effort-reward imbalance at work: European comparisons. Soc Sci Med 2004;58(8):1483-1499. [Medline: 14759692]
42. Siegrist J, Li J, Montano D. University Hospital Dusseldorf. 2014 Jun 4. Psychometric properties of the Effort-Reward Imbalance Questionnaire URL: https://www.uniklinik-duesseldorf.de/fileadmin/Fuer-Patienten-und-Besucher/Kliniken-Zentren-Institute/Institute/Institut_fuer_Medizinische_Soziologie/Forschung/PsychometricProperties.pdf [accessed 2019-06-04] [WebCite Cache ID 78taRWiKQ]
43. Miller WL, Crabtree BF, Harrison MI, Fennell ML. Integrating mixed methods in health services and delivery system research. Health Serv Res 2013 Dec;48(6 Pt 2):2125-2133 [FREE Full text] [doi: 10.1111/1475-6773.12123] [Medline: 24279834]
44. Sale JE, Lohfeld LH, Brazil K. Revisiting the quantitative-qualitative debate: implications for mixed-methods research. Qual Quant 2002 Feb;36(1):43-53 [FREE Full text] [doi: 10.1023/A:1014301607592] [Medline: 26523073]
45. R Core Team. R: A language and environment for statistical computing. 2017. URL: https://www.R-project.org/ [WebCite Cache ID 72QzjOUtv]
46. Creswell J. Qualitative Inquiry & Research Design: Choosing among Five Approaches. Los Angeles: Sage Publications; 2012.
47. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]

Abbreviations

DCSQ: Demand-Control-Support Questionnaire
ERI: effort-reward-imbalance
MAAS: Mindful Attention Awareness Scale
MBI: mindfulness-based intervention
PTSD: posttraumatic stress disorder
VAS: visual analog scale

Edited by G Eysenbach; submitted 20.01.19; peer-reviewed by C Young, MS Aslam; comments to author 14.02.19; revised version received 09.04.19; accepted 05.05.19; published 19.06.19.

Please cite as:
Kerr DC, Ornelas IJ, Lilly MM, Calhoun R, Meischke H
Participant Engagement in and Perspectives on a Web-Based Mindfulness Intervention for 9-1-1 Telecommunicators: Multimethod Study
J Med Internet Res 2019;21(6):e13449
URL: http://www.jmir.org/2019/6/e13449/
doi:10.2196/13449
PMID:31219045

©Darragh C Kerr, India J Ornelas, Michelle M Lilly, Rebecca Calhoun, Hendrika Meischke. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Persuasive System Design Principles and Behavior Change Techniques to Stimulate Motivation and Adherence in Electronic Health Interventions to Support Weight Loss Maintenance: Scoping Review

Rikke Aune Asbjørnsen[1,2,3], MSc; Mirjam Lien Smedsrød[4], MA; Lise Solberg Nes[3,5,6], PhD; Jobke Wentzel[1,7], PhD; Cecilie Varsi[3], PhD; Jøran Hjelmesæth[8,9], PhD; Julia EWC van Gemert-Pijnen[1,10,11], PhD

[1]Center for eHealth and Wellbeing Research, Department of Psychology, Health, and Technology, University of Twente, Enschede, Netherlands
[2]Research and Innovation Department, Vestfold Hospital Trust, Tønsberg, Norway
[3]Center for Shared Decision Making and Collaborative Care Research, Division of Medicine, Oslo University Hospital, Oslo, Norway
[4]Norwegian Regional Advisory Unit on Patient Education, Sørlandet Hospital Trust, Kristiansand, Norway
[5]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
[6]Department of Psychiatry & Psychology, Mayo Clinic, Rochester, MN, United States
[7]Saxion University of Applied Sciences, Deventer, Netherlands
[8]Morbid Obesity Center, Vestfold Hospital Trust, Tønsberg, Norway
[9]Department of Endocrinology, Morbid Obesity, and Preventive Medicine, Institute of Clinical Medicine, University of Oslo, Oslo, Norway
[10]University Medical Center Groningen, Groningen, Netherlands
[11]University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Rikke Aune Asbjørnsen, MSc
Center for eHealth and Wellbeing Research
Department of Psychology, Health, and Technology
University of Twente
De Zul 10
Enschede, 7522 NJ
Netherlands
Phone: 31 534899111
Email: r.a.asbjornsen@utwente.nl

Abstract

Background: Maintaining weight after weight loss is a major health challenge, and eHealth (electronic health) solutions may be a way to meet this challenge. Application of behavior change techniques (BCTs) and persuasive system design (PSD) principles in eHealth development may contribute to the design of technologies that positively influence behavior and motivation to support the sustainable health behavior change needed.

Objective: This review aimed to identify BCTs and PSD principles applied in eHealth interventions to support weight loss and weight loss maintenance, as well as techniques and principles applied to stimulate *motivation* and *adherence* for long-term weight loss maintenance.

Methods: A systematic literature search was conducted in PsycINFO, Ovid MEDLINE (including PubMed), EMBASE, Scopus, Web of Science, and AMED, from January 1, 2007 to June 30, 2018. Arksey and O'Malley's scoping review methodology was applied. Publications on eHealth interventions were included if focusing on weight loss or weight loss maintenance, in combination with motivation or adherence and behavior change.

Results: The search identified 317 publications, of which 45 met the inclusion criteria. Of the 45 publications, 11 (24%) focused on weight loss maintenance, and 34 (76%) focused on weight loss. Mobile phones were the most frequently used technology (28/45, 62%). Frequently used wearables were activity trackers (14/45, 31%), as well as other monitoring technologies such as wireless or digital scales (8/45, 18%). All included publications were anchored in behavior change theories. *Feedback and monitoring* and *goals and planning* were core behavior change technique clusters applied in the majority of included publications.

XSL•FO
RenderX

Social support and *associations* through prompts and cues to support and maintain new habits were more frequently used in weight loss maintenance than weight loss interventions. In both types of interventions, frequently applied persuasive principles were *self-monitoring*, *goal setting*, and *feedback*. *Tailoring*, *reminders*, *personalization*, and *rewards* were additional principles frequently applied in weight loss maintenance interventions. Results did not reveal an *ideal* combination of techniques or principles to stimulate motivation, adherence, and weight loss maintenance. However, the most frequently mentioned individual techniques and principles applied to stimulate motivation were, *personalization*, *simulation*, *praise*, and *feedback*, whereas *associations* were frequently mentioned to stimulate adherence. eHealth interventions that found significant effects for weight loss maintenance all applied *self-monitoring*, *feedback*, *goal setting*, and *shaping knowledge*, combined with a human *social support* component to support healthy behaviors.

Conclusions: To our knowledge, this is the first review examining key BCTs and PSD principles applied in *weight loss maintenance* interventions compared with those of *weight loss* interventions. This review identified several techniques and principles applied to stimulate motivation and adherence. Future research should aim to examine which eHealth design combinations can be the most effective in support of long-term behavior change and weight loss maintenance.

(J Med Internet Res 2019;21(6):e14265) doi:10.2196/14265

KEYWORDS

eHealth; weight loss maintenance; weight loss; behavior change; persuasive technology; review; motivation; adherence

Introduction

The Weight Loss Maintenance Challenge

Obesity is a rapidly increasing public health problem, with more than 600 million people with obesity (body mass index [BMI] ≥30 kg/m²) worldwide [1,2]. One of the main challenges related to obesity is that although many people manage to lose weight, only 1 of 4 people manages to maintain the lost weight in the long term [3]. With several mechanisms interacting (eg, biological, environmental, behavioral, and cognitive) to make weight regain following weight loss common [4,5], novel ways to counterbalance these mechanisms are needed [6,7]. The significant weight loss maintenance challenge calls for the examination of new technologies and solutions in future studies of long-term weight control [6,8,9].

Electronic Health Design for Sustainable Behavior Change

eHealth (electronic health) is a term often used to define the blending of digital technologies, health, and health services and can be delivered through information and communication technologies [10-12]. Although evidence is sparse regarding the impact of eHealth on health and health care, research indicates that eHealth can support health care delivery by providing greater efficiency, lead to better health outcomes, and lower health service costs [10,13-15]. eHealth technologies are also increasingly used to support a healthier life, improved well-being, and creation of new health behaviors [16-19] and have the potential to support the challenging behavior changes needed to sustain long-term weight loss maintenance [4,20].

Behavior change interventions are usually complex and may include many interacting components or techniques [21,22]. Behavior change techniques (BCTs) are observable and active intervention components aiming to regulate behavior (eg, goal setting, self-monitoring, and feedback) [21,23]. However, health behavior change requires motivation and persistence, and persuasive design [24] also appears to play an important role in this setting. Persuasive design of services or technologies focuses on influencing human behavior in a positive way. As such, persuasive system design (PSD) principles can be applied in eHealth design to match user profiles, motivate users to engage in self-management, and trigger health behavior change [16,24-26]. Several behavior change theories, BCTs, and PSD principles can be involved in an eHealth intervention [27], alone or in combination.

To date, there is limited knowledge about how behavior change interventions and design of technologies and services can impact behavior and motivation in support of sustainable health behavior change [25,28,29]. eHealth is often described as a *black box*, as knowledge is limited about its internal structure and how the use of various components of the technology can contribute to healthier lifestyles and improved health outcomes [11,22,30,31]. Finding the right mix of technological features to stimulate the motivation and adherence needed to support long-term weight loss maintenance is, therefore, still a conundrum [7]. Little is also known about how BCTs and PSD principles can be used in eHealth interventions to support long-term weight loss maintenance [8,32]. The application of the most effective BCTs and PSD principles, at the right time and in the best combination, could therefore be of essence to support motivation and adherence in the pursuit of sustainable weight loss maintenance [19,33-35].

The Goal of This Review

The overall goal of this review was to provide insight into the design of eHealth interventions aiming to support behavior change for long-term weight loss maintenance in adult people with obesity. This review identified BCTs and PSD principles to stimulate motivation and adherence in eHealth interventions built to support weight loss maintenance.

Research questions for this review are as follows: in eHealth interventions, (1) how are *motivation* and *adherence* defined and measured? (2) Are *motivation* and *adherence* linked to weight loss and weight loss maintenance? (3) What can be determined from behavior change theories, BCTs, and PSD principles used in weight loss and weight loss maintenance interventions? (4) Which behavior change theories, BCTs, and PSD principles have been used to stimulate motivation and

XSL•FO
RenderX

adherence in eHealth weight loss maintenance interventions, and in what combination? (5) What are the reported effects (ie, weight outcomes) in eHealth weight loss maintenance interventions?

Methods

Scoping Review Methodology

A scoping review methodology was considered suitable for mapping literature on BCTs and PSD principles, as this is an emerging topic where evidence is scarce and key concepts and gaps in existing research should be identified [36,37]. This scoping review applied the methodology by Arksey and O'Malley [36], with the following steps [37]: (1) identify the research questions; (2) identify relevant studies; (3) study selection; (4) chart the data; (5) collate, summarize, and report the results; and (6) consultation. To enhance the scoping study methodology, additional recommendations [37] were followed: (1) 2 independent researchers reviewed all full-text publications, and (2) the research group developed and continuously updated the data extraction form during the extraction process.

As research on eHealth interventions targeting weight loss maintenance is still in its infancy, eHealth interventions targeting weight loss were also examined to best identify weight loss maintenance–related factors. Research questions 1, 2, and 3 entailed broad scopes. The scope was then further narrowed in research questions 4 and 5, focusing on weight loss maintenance interventions and the BCTs and PSD principles applied to stimulate motivation and adherence, as well as any effects (ie, weight outcomes) related to weight loss maintenance.

This review applied Michie's Behavior Change Taxonomy [21] developed to meet the need for standardized reporting on development and content of complex behavior interventions. Michie's Cross-Domain Taxonomy consists of 93 distinct BCTs divided into 16 clusters, independent of any specific theory. Multimedia Appendix 1 provides detailed information about the BCT clusters. Similarly, the PSD model by Oinas-Kukkonen [16], building on previous research by Fogg [26], was used as a framework to identify persuasive principles applied in the included interventions. Multimedia Appendix 2 provides information about the 4 PSD categories: *primary task support*, *dialog support*, *system credibility support*, and *social support*, as well as operationalization of the individual principles [16]. For the purpose of this review, adherence to a technology was defined as *use as intended or desired* by the authors or developers of an intervention [19], whereas motivation was defined as *a reason for doing something* [38].

Search Strategy

A systematic literature search to cover behavioral, technical, and clinical research aspects was conducted in the following databases: PsycINFO, Ovid MEDLINE (including PubMed), EMBASE, Scopus, Web of Science, and AMED. As digital technologies are advancing and developing fast, more recent evidence (ie, since 2007) was considered to be the most relevant and interesting. Publications during the period from January 1, 2007 to June 30, 2018, were therefore included. The terms *weight loss* and *weight loss maintenance* were used, in combination with a variety of the term *eHealth interventions* and the terms *motivation*, *adherence*, and/or *behavior change*. This search strategy was created and applied in close collaboration with librarians and domain experts (Multimedia Appendix 3).

Eligibility Criteria

Publications in English, clearly describing an eHealth intervention focusing on weight loss maintenance or weight loss, were included for assessment when containing persuasive design, behavior change theories, and techniques or when mentioning motivation and/or adherence. The target population for this review was people with overweight (ie, BMI 25-29.9 kg/m^2) and/or obesity (ie, BMI ≥30 kg/m^2). Multimedia Appendix 4 gives a complete overview of the inclusion and exclusion criteria.

Data Collection and Analysis

A data charting form containing general as well as specific study characteristics was created in Microsoft Excel by the research team. The characteristics were extracted using elements from the CONSORT-eHealth checklist [39], focusing on characteristics about the interventions and technologies in the included publications. Michie's Cross-Behavior Change Taxonomy [21] guided the extraction process to identify and group specific information about BCTs used. For the purpose of this review, representation of the 16 clusters as indicated in Multimedia Appendix 1 was applied, rather than presenting a detailed representation of up to 93 distinct techniques. Supplementary information about definitions, including specific examples, was reviewed [40], and behavior change theories mentioned or described were recorded. Persuasive principles were extracted from the included publications using the PSD model [16] presented in Multimedia Appendix 2. PSD principles were coded when executed by or through the technology. Due to lack of reporting on the system credibility category, this category was not part of the analysis. *Goal setting*, *feedback*, and *social support* were added to the model as separate persuasive principles, as they could not always easily be linked to specific design elements in the PSD model.

Two researchers (RA and MS) independently coded and categorized the identified BCTs and PSD principles, using the presented frameworks. A third researcher (CV) validated 11% (5/45) of the included interventions. The first and second author also recorded additional information on motivation and adherence when mentioned, including how motivation/adherence was defined, stimulated, and measured. Data on effect (ie, weight outcomes) were recorded when reported, including when weight was self-reported or measured by the researcher or coach/clinician. The research team also extracted relevant information (eg, intervention components, BCTs, and PSD principles) from the incorporated interventions, including illustrations, figures, tables, and additional websites when referred to in the publications. To enhance and support the relevance of the review, clinicians and researchers specialized in the fields of weight loss/maintenance, health psychology, and eHealth were consulted regarding methodological approach, relevance, and current state of the evidence.

XSL•FO
RenderX

Weight Loss and Weight Loss Maintenance

Weight loss BCTs do not necessarily equal weight loss maintenance BCTs. Long-term maintenance of lost weight is challenging, and there is a call for interventions evaluating novel methods to improve the maintenance of lost weight [8,41]. To meet this call and the overall goal of this review (ie, supporting long-term weight loss maintenance), the first part of the Results section focuses on weight loss and weight loss maintenance, whereas the rest of the results section focuses solely on weight loss maintenance interventions.

Results

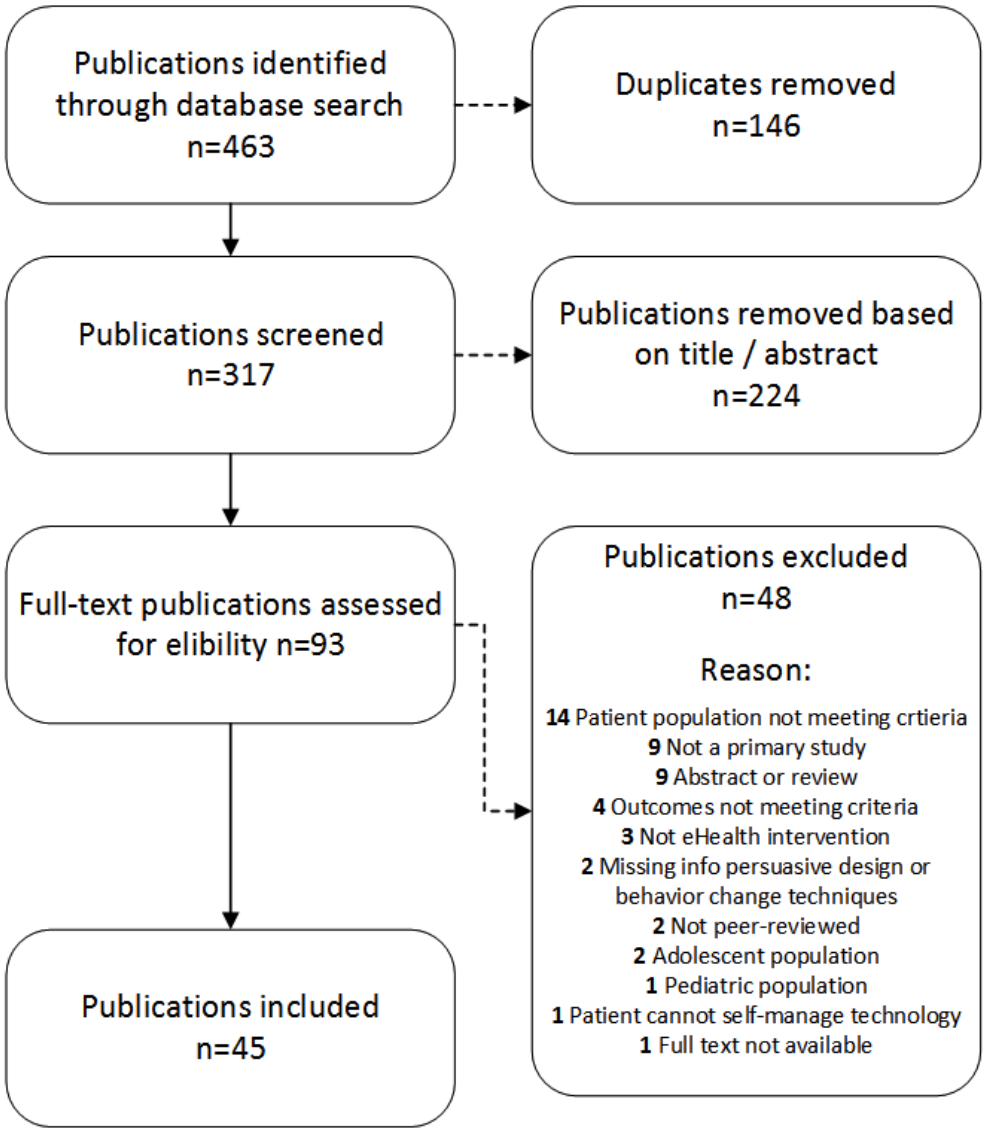
Study Selection

The search revealed 463 publications. Following removal of 146 duplicates, the remaining 317 titles and abstracts were screened for eligibility by 2 researchers (RA and either MS, CV or FS) using the Covidence software program [42]. After removing 224 publications not meeting eligibility based on title/abstracts, the remaining 93 full-text publications were reviewed by the first (RA) and second (MS) authors. The authors RA and MS discussed differences and disagreements until consensus was reached. If consensus could not be reached, consultation was sought from a third researcher (FS). After the full-text screening process, 45 publications remained to be assessed for further analyses (Figure 1).

General Characteristics of Included Publications

All 45 publications included in this review were published in 2008 or later (Figure 2). First authors were mainly affiliated in the United States (25/45, 56%), the United Kingdom (5/45, 11%), and Australia (3/45, 7%); see Table 1.

Figure 1. Flow diagram of study selection process. n=the actual number of publications. eHealth: electronic health.



XSL•FO
RenderX

Figure 2. The number of included publications per year categorized by the aim of the electronic health intervention.

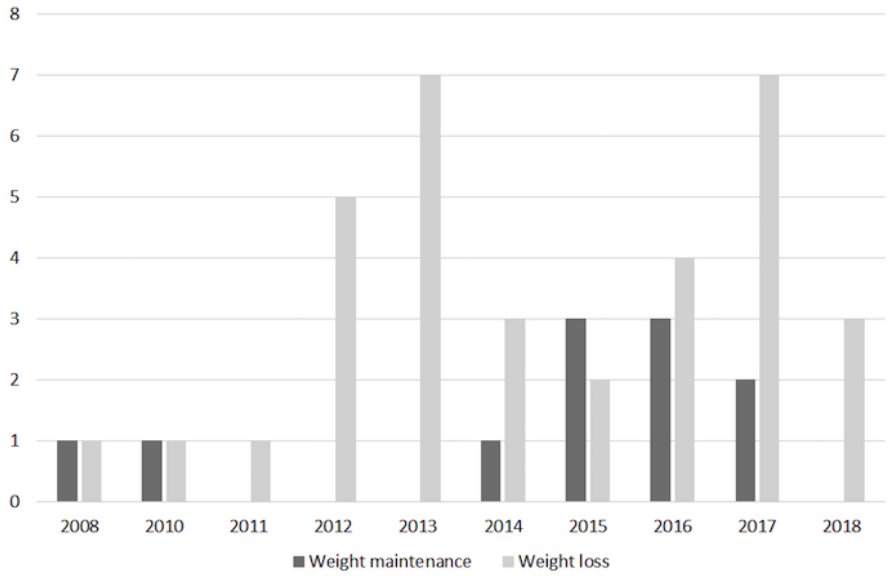


Table 1. Country of affiliation for the first authors of all included publications (N=45).

Country of origin	Included publications, n (%[a])
United States	25 (56)
United Kingdom	5 (11)
Australia	3 (7)
Canada	2 (4)
The Netherlands	2 (4)
Finland	1 (2)
Germany	1 (2)
Italy	1 (2)
Scotland	1 (2)
Saudi Arabia	1 (2)
Spain	1 (2)
Romania	1 (2)
Qatar	1 (2)

[a]Percentages do not total 100% due to rounding errors.

Of the 45 included interventions, 34 (76%) targeted weight loss, and 11 (24%) targeted weight loss maintenance. Of 11 weight loss maintenance interventions, 5 (45%) had an initial weight loss phase. The most frequently described study design was randomized controlled trials for the weight loss (15/45, 33%) and weight loss maintenance (4/45, 9%) interventions alike. Although several of the included interventions evaluated effects (ie, weight outcome; 26/45, 58%), others described only the design of the intervention (6/45, 13%) or a protocol (7/45, 16%). Most interventions targeted people with overweight (24/45, 53%) or included both overweight and obesity (15/45, 33%). People with obesity (ie, BMI ≥30 kg/m^2) were the sole target population in only 4 weight loss interventions and 2 weight loss maintenance interventions. Average duration of interventions was 24 weeks (range 4-104; median 13 weeks) for the weight loss interventions, and 27 weeks (range 12-52; median 26 weeks) for the weight loss maintenance interventions. User involvement during the technology development process was only described by 8 weight loss and 4 weight loss maintenance interventions. The users were usually only mentioned when involved in part of the development process (eg, identifying needs, content development, and usability testing). Two publications were included despite describing the same intervention, as they focused on different aspects of the intervention [43,44]. Multimedia Appendix 5 provides an overview of the included interventions including title, authors, publication year, country of origin, design, objectives, participants, aim and type of technology, intervention duration, and whether blended care was part of the intervention or not.

XSL•FO
RenderX

Technology Characteristics

In the 45 included publications, mobile phones were the most frequently used technology (28/45, 62%), followed by Web-based solutions through computers (15/45, 33%), or a combination of computer and mobile phone (eg, for feedback or reminders through text messages; 6/45, 13%). Monitoring technologies used were activity trackers or step counters (14/45, 31%), wireless or digital weight scales (8/45, 18%), and glucose (2/45, 4%) or blood pressure (1/45, 2%) monitors, often combined with manually recorded self-monitoring data. Some technologies included game-based elements (3/45, 7%) [45-47], a virtual world with avatars [48] or a virtual coach [49], and other tools to enhance self-monitoring (eg, automated calculations of energy intake, expenditure, and energy balance). Several eHealth weight loss interventions (20/45, 44%) and weight loss maintenance interventions (7/45, 16%) had a blended care approach, including the combination of various formats of human coaching and/or face-to-face services from experts. Some eHealth interventions also integrated peer forums or social media groups (7/45, 16%) or included social support through a buddy or helper (eg, family, friend, colleague; 3/45, 7%).

The typical eHealth weight loss maintenance intervention was supported by mobile phone technology (9/11, 82%), in combination with an activity tracker or step counter (3/11, 27%) and/or wireless scale (3/11, 27%) [47,50-52]. The technology usually supported 2-way communication with a peer, dietitian, or coach [50,53-56] and provided automated, tailored feedback based on progress data [50-53,55,57,58]. The technology in these weight loss maintenance interventions usually aimed to support creation of healthy habits [51-54]; educational resources and information [51,53,54,59]; daily or weekly monitoring tools for weight, diet, activity (eg, number of steps) [47,50-52,55]; well-being (eg, mood, stress, and good days/bad days) [47,51,53]; and/or plans or strategies for individual action and coping [47,54,57].

Motivation Defined, Measured, and Linked to Weight Loss and Weight Loss Maintenance

Only 2 of the 45 included publications provided a *definition* of motivation [43,44], and these 2 publications originated from the same intervention. By referring to self-determination theory, the publications distinguished between autonomous motivation (a measure of a person's *internal or personal reasons for change*) or controlled motivation (a measure of *extrinsic reasons or external pressure to change*) [44].

Motivation was *measured* in various ways by self-reported measures and questionnaires as presented in Table 2.

Motivation was *evaluated* (ie, *link to weight loss/maintenance*) in 7 of 35 weight loss intervention studies [43,44,49,60,63-65], but in none of the included weight loss maintenance interventions (n=11) as indicated in Multimedia Appendix 6. Of the 7 weight loss studies that evaluated motivation, 1 found high levels of controlled motivation at baseline to produce significantly greater weight loss in the motivation-enhanced intervention (ie, specific components were used to enhance autonomous motivation) compared with the standard intervention [43]. The motivation-enhanced group used the website more often than the control group, and the number of visits was associated with weight loss [43]. To increase autonomous motivation, principles of motivational interviewing [66] together with goal setting and journaling (eg, writing about the future when weight loss goals are achieved) and blended formats, such as face-to-face sessions, were also added to the Web-based weight loss program to improve autonomous motivation [43].

Another study showed the level of autonomous motivation after 4 weeks to be predictive of self-monitoring of adherence and weight loss at 16-week postbaseline [44]. For participants who reached 5% weight loss, autonomous motivation increased and remained higher than for those not reaching this clinically meaningful weight loss [44]. A third study suggested that diet-focused constructs were particularly important when developing weight loss interventions for men [60] because changes in diet-related autonomous motivation were linked to weight loss. In that study, the intervention group achieved greater weight loss than the control group [60].

The fourth study evaluating motivation found motivational orientation (eg, promotion focused or prevention focused) to be a predictor of behavior change when trying to lose weight, and framing messages with people's motivational orientation was considered preferable to, for example, informational and prescriptive messages in terms of behavior change [63]. As high as 72% of the participants in another study found text messages received biweekly to be motivational [64], and 79% reported text messages to be helpful in performing healthy eating and exercise behaviors [64]. The sixth study evaluating motivation found that delivery of remote daily real-time feedback messages tailored to diary entries could enhance motivation, producing greater reductions in energy and saturated fat consumptions [67]. In the final study, 58% of the participants agreed that a virtual coach motivated them to become more active [49], suggesting that meetings with a virtual coach could be beneficial in maintaining activity level. However, no significant changes in step count were found in the intervention versus the control groups [49].

XSL•FO
RenderX

Table 2. Measurements of motivation.

Methods for measuring motivation	Reasons to measure motivation	Publications[a]	P[b]/I[c]
Treatment Self-Regulation Questionnaire	To assess autonomous versus controlled motivation for self-regulation, weight loss, healthy eating, and continued exercise	WL9[d] [60]; WL25 [61]; WL30 [43]; WL31[44];	I; P; I; I
Behavioral Regulation in Physical Exercise and Eating Habit Questionnaire	To assess treatment moderators (here, intrinsic motivation) and measure stages of the self-determination continuum, a motivation factor	WL4 [62]	P
Consideration of Future Consequences	To characterize motivational orientation and measure participants' focus on distal versus proximal consequences/outcomes of behaviors	WL8 [63]	I
Behavioral Inhibition and Behavioral Activation Scales	To identify participants' motivational orientation, either predominantly promotion focused (gain focus) or predominantly prevention focused (loss focus)	WL8 [63]	I
The Diet and Exercise Self-Efficacy Questionnaires	To assess self-efficacy to make and maintain diet and exercise behavior changes	WL8 [63]	I
The University of Rhode Island Change Assessment scale	On the basis of the transtheoretical (stages-of-change) model, to assess where an individual exists along a 5-phase continuum from precontemplation to contemplation, preparation, action, and maintenance	WL8 [63]	I
Online survey	To map to what extent text messages were experienced as motivational	WL5 [64]	I
Online self-reporting	To rate their motivation and confidence to continue their weight next week	WM2[e] [51]	P
Online self-reporting/feedback	To set the level of participant motivation	WL23 [65]	I

[a]Multimedia Appendix 5 provides an overview of the publications.

[b]P: protocol.

[c]I: intervention.

[d]WL: weight loss.

[e]WM: weight loss maintenance.

Adherence Defined, Measured, and Linked to Weight Loss and Weight Loss Maintenance

Of the 45 included publications in this review, 6 weight loss interventions measuring adherence provided a *definition* of the adherence concept or related terms such as usage and compliance [64,68-72]. These publications defined or operationalized adherence based on either self-monitoring or electronic entry of food and exercise records [68]. Low usage was defined as having no food records and high usage as having 1 food record on a randomly selected day of the sampled week [68] or as recording of 50% or more of prescribed daily calorie intake goal [69]. Some publications also defined adherence as responsiveness to text messages or health challenges [64,71,72] or looked at various aspects of adherence such as behavioral adherence (ie, attendance to counseling sessions) [68] and dietary adherence (ie, self-monitoring related to dietary goals) [68,70]. Other publications defined adherence as program compliance to habits and workout [71] or consistency to self-monitoring [73].

Regarding *measurement* of adherence, as the main scope of this review was weight loss maintenance, the results related to short-term weight loss interventions (≤6 months) measuring adherence are not reported in this review [18,43,44,49,69,70,73-77]. Moreover, 4 of the included long-term weight loss interventions (>6 months) measured adherence [68,71,72,78]. For weight loss maintenance, adherence was measured in 3 of the interventions [50,52,59], and the duration for 2 of these interventions was more than 6 months (Multimedia Appendix 6). The 4 long-term weight loss interventions measured various aspects of adherence, including self-monitoring data related to diet and physical activity [68], compliance to the Web-based program, daily habits and exercise [71], website usage by the number of self-tracking entries [78], or the total percentage of text messages that a participant responded to [72]. The 3 weight loss maintenance interventions measured adherence or engagement in relation to the coaching program through evaluating frequency of submitting self-monitoring data to their coach [52], the number of delivered text messages replied to by the participants [59], and by participants' self-monitoring adherence through frequency of weigh-ins and use of activity tracking [50].

Adherence to technology was defined as *use as intended or desired* in this review [19]. Intended usage was reported in the 4 weight loss interventions (>6 months) and in all 3 weight loss maintenance interventions. Intended use was most frequently described as 1 time per day or more (≥1 per day).

Links to Weight Loss and Weight Loss Maintenance

Although actual technology usage was not evaluated in the 4 weight loss interventions measuring adherence, and significant results were sparsely reported, some interventions measured certain aspects of the technology features and linked these to weight loss. However, 1 intervention showed that participants with a high usage of self-tracking entries initially lost greater amounts of weight than participants with low usage [68]. Other

XSL•FO
RenderX

interventions reported that compliance to the Web-based program, daily habits, and exercise was also found to be a significant predictor of weight loss [71], and that participants with greater adherence to text messages lost more weight [72].

The 3 weight loss maintenance interventions measured adherence, showing significant effects (ie, weight loss maintenance) at 12 weeks postbaseline, 12 months postintervention, and 24 months postbaseline [50,52,59]. It should be noted that the methods applied and/or results reported only provided results related to intervention engagement and technology or intervention features, not related to the actual use of the technology [50,52,59] (Multimedia Appendix 6).

Behavior Change Theories and Behavior Change Techniques in Weight Loss and Weight Loss Maintenance Interventions

Behavior Change Theories Applied in Weight Loss and Weight Loss Maintenance Interventions

All 45 included publications were theoretically anchored in behavior change theories. In approximately two-thirds of the interventions, specific behavioral change theories were mentioned as applied. Interventions that did not specify behavior change theories referred to behavioral strategies or techniques as crucial factors for behavior change [18,50,52,55,58,62,64,67-70,75,78-82]. Multimedia Appendix 7 shows an overview of the behavior change theories specified as applied in the weight loss and weight loss maintenance interventions.

Behavior change technique clusters applied in weight loss and weight loss maintenance interventions

Analysis of all included publications (N=45) identified 15 of the 16 BCT clusters specified in Michie's taxonomy [21] (Table 3).

The *goal and planning* and *feedback and monitoring* clusters were referred to as core self-regulation techniques for behavior change and weight outcomes in several of the publications [51,52,54,74,77]. These cluster techniques were also applied in most of the weight loss *and* weight loss maintenance interventions. Techniques contributing to the cluster *shaping knowledge* were present in 82% of the interventions (weight loss and weight loss maintenance). This cluster included providing relevant information on diet, physical activity, and how to change behavior, advice on how to perform a desired behavior, or advice to keep a record on social situations, emotions, or cognitions that typically occur before temptations (eg, snacking) [47,50,52,54,56,58,59]. *Social support* was a more frequently used technique in weight loss maintenance (91%) than in weight loss (68%) interventions, enabled with as well as without technology. *Social support* was typically provided through e-coaching and social reinforcement from professionals or peers [45,51-53,58,80], with encouragement and counseling on performed behavior [52,57].

Table 3. Behavior change cluster of techniques according to Michie's taxonomy.

Cluster labels	WM[a] (n=11), n (%)	WL[b] (n=34), n (%)	All (N=45), n (%)
Scheduled consequences	1 (9)	2 (6)	3 (7)
Reward and threat	3 (27)	4 (12)	7 (16)
Repetition and substitution	8 (73)	24 (71)	32 (71)
Antecedents	4 (36)	8 (24)	12 (27)
Associations	8 (73)	15 (44)	23 (51)
Covert learning	0 (0)	0 (0)	0 (0)
Natural consequences	3 (27)	7 (21)	10 (22)
Feedback and monitoring	11 (100)	34 (100)	45 (100)
Goals and planning	11 (100)	33 (97)	44 (98)
Social support	10 (91)	23 (68)	33 (73)
Comparison of behavior	0 (0)	9 (26)	9 (20)
Self-belief	2 (18)	6 (18)	8 (18)
Comparison of outcomes	2 (18)	4 (12)	6 (13)
Identity	1 (9)	4 (12)	5 (11)
Shaping knowledge	9 (82)	28 (82)	37 (82)
Regulations	3 (27)	3 (9)	6 (13)

[a]WM: weight loss maintenance interventions.

[b]WL: weight loss interventions (Multimedia Appendix 5 provides an overview of the publications).

Michie's cluster *associations* were more frequently applied in weight loss maintenance (73%) than in weight loss (44%) interventions, and these techniques were often an environmental or social stimulus or reminders with the purpose of prompting a specific behavior [52,54,56,57]. The *comparison of behavior* cluster was only present in 26% of the weight loss interventions

XSL•FO
RenderX

and not identified at all in the weight loss maintenance interventions [43-45,71,75,80,83,84].

Persuasive System Design Principles Applied in Weight Loss and Weight Loss Maintenance Interventions

An overview of the PSD principles applied by or through the technology in the included publications (N=45) is presented in Table 4. In the included interventions, the *primary task support* category from the PSD model [16], supporting users to do primary tasks, was applied most often (50%), followed by *dialog support* (35%) and *social support* (15%).

Table 4. Persuasive system design principles.

Persuasive principles	WM[a] (n=11), n (%)	WL[b] (n=34), n (%)	All (N=45), n (%)
Primary task support			
Self-monitoring	11 (100)	30 (88)	41 (91)
Tailoring	11 (100)	22 (65)	33 (73)
Personalization	8 (73)	16 (47)	24 (53)
Simulation	8 (73)	15 (44)	23 (51)
Reduction	3 (27)	4 (12)	7 (16)
Tunneling	3 (27)	5 (15)	8 (18)
Rehearsal	2 (18)	5 (15)	7 (16)
Dialog support			
Reminders	9 (82)	15 (44)	24 (53)
Suggestions	7 (64)	20 (59)	27 (60)
Reward	6 (55)	5 (15)	11 (24)
Praise	4 (36)	13 (38)	17 (38)
Social role	1 (9)	2 (6)	3 (7)
Similarity	0 (0)	2 (6)	2 (4)
Liking	0 (0)	1 (3)	1 (2)
Social support			
Social comparison	2 (18)	8 (24)	10 (22)
Social facilitation	2 (18)	3 (9)	5 (11)
Social learning	1 (9)	7 (21)	8 (18)
Cooperation	1 (9)	3 (9)	4 (9)
Recognition	1 (9)	3 (9)	4 (9)
Competition	0 (0)	4 (12)	4 (9)
Normative influence	0 (0)	1 (3)	1 (2)
Other			
Feedback	11 (100)	31 (91)	42 (93)
Goal setting	10 (91)	27 (79)	37 (82)
Social support	7 (64)	18 (53)	25 (56)

[a]WM: weight loss maintenance interventions.

[b]WL: weight loss interventions.

In the weight loss interventions, the most frequently applied persuasive principles were *feedback* (91%), *self-monitoring* (88%), *goal setting* (79%), *tailoring* (65%), and *suggestions* (59%). In the weight loss maintenance interventions, *feedback*, *self-monitoring*, and *tailoring* (all 100%) were the most frequently applied persuasive principles, followed by *goal setting* (91%) and *reminders* (82%). *Social support* as a PSD principle, usually 2-way communication with peers or a coach, was used to support continued behavior change [46,54,56,84] and was identified and applied almost as often in weight loss (53%) as in weight loss maintenance (64%) interventions. Frequently reported PSD principles in weight loss maintenance interventions compared with weight loss interventions were *personalization* (73% vs 47%), *simulation* (73% vs 44%), and *rewards* (55% vs 15%). *Competition*, on the other hand, was one of the persuasive principles not identified in any weight

XSL•FO
RenderX

loss maintenance interventions, although identified in 12% of the weight loss interventions, often related to weight changes and/or activity targets [45,75,84,85]. Application of the *self-monitoring* principle, one of the most frequently applied principles in both types of interventions, was associated with user recording of weight and behaviors connected to diet and physical activity targets [51,53,57] and reception of automated, *tailored feedback* through text messages or visually by graphs, charts, bars, symbols (eg, traffic light, colors when entering a *danger zone*) [51], and dashboard [50] related to their progress. *Normative influence*, *similarity*, and *liking* were the persuasive principles least applied in both types of interventions (Table 4). An overview of PSD principles identified in the included weight loss and weight loss maintenance interventions can be found in Multimedia Appendix 8.

Behavior Change Theories, Behavior Change Techniques, and Persuasive System Design Principles Used to Stimulate Motivation and Adherence in Electronic Health Weight Loss Maintenance Interventions

As seen in Tables 3 and 4, weight loss BCTs and PSD principles do not necessarily equal weight loss maintenance techniques. To meet the call for interventions evaluating novel methods to improve maintenance of lost weight [8] and meet the overall goal of this review, the next part of the Results section focuses solely on weight loss maintenance interventions.

Behavior Change Theories and Techniques Used to Stimulate Motivation and Adherence in Electronic Health Weight Loss Maintenance Interventions

Behavior Change Theories

Of the 11 weight loss maintenance interventions included in this review, 7 explicitly mentioned which behavior change theories were used [47,51,53,54,56,57,59] but did not describe using these to stimulate motivation or adherence directly as indicated in Table 5. Of the publications that described reasons for applying the identified behavioral theories, the following were mentioned: (1) to support long-term behavior change by increasing the individual coping capacity [57], (2) to facilitate goal setting, monitoring, and feedback [51], (3) to adapt the text messages to participant's readiness for change [59], (4) to support existing behavior change, and (5) to develop new self-management skills [56]. Motivational interviewing was added to 1 eHealth intervention [53] through physical consultation to enhance adherence.

Behavior Change Techniques

Table 5 shows that all 11 weight loss maintenance interventions used various BCTs to stimulate weight loss maintenance. Of 11 interventions, 9 applied BCT clusters to stimulate motivation and/or adherence. Analyses indicate that *feedback and monitoring* was the most frequently mentioned cluster to stimulate motivation (55%) [47,51,53,54,57,58]. Within this cluster, motivation was enhanced through encouraging and supporting usually by automated messages [47,51,53,57]. Face-to-face contact was also offered to enhance participants' motivation to engage with internet-delivered elements [51]. *Monitoring* was also mentioned as a key strategy to achieve weight behavior change and weight control in several studies, and different BCTs were used to support manually and automated *monitoring* of goals and target behavior [47,50-54,56,57]. *Associations* were, as indicated in Table 5, applied most often to stimulate adherence through *prompts and cues* in the weight loss maintenance interventions (27%) [54,56,57].

In 1 study, 3 clusters of BCTs, *feedback and monitoring*, *goals and planning*, and *social support*, were applied to stimulate motivation [52]. Target behavior was emphasized to be reduced from great overall goals to *small steps*, which could be more easily reached by the participants [53]. Other studies showed *reward* as 1 of the clusters mentioned applied to stimulate motivation as well as adherence in support of weight loss maintenance by linking financial rewards to submitting self-monitoring records ($1–$10 per week) [52] or accomplishing behavioral goals (eg, by offering direct payments ($2.80) each day the participants weighed in and met the weight loss goals) [55]. Incentive *payouts* were also made contingent on a *dyadic partner performance* to stimulate adherence [52], meaning that members of the partner dyad had to email for 5 days or more about self-monitoring, and both partners had to maintain their weight loss to get the financial incentives. The cluster *shaping knowledge* was used to offer general suggestions and theory-based advice on how to maintain weight loss, not to stimulate motivation or adherence (eg, information and supportive tools on diet, physical activity, and behavior change) [50,51,54,56,59]. A serious game-based eHealth intervention also offered access to information while playing (eg, general information about dieting, research news, and fact sheets) [58].

XSL•FO
RenderX

Table 5. Included weight loss maintenance interventions specifying behavior change theories and behavior change technique (BCT) clusters.

Study ID[a]	WM[b] 1 [57]	WM 2 [51]	WM 3 [52]	WM 4 [59]	WM 5 [55]	WM 6 [56]	WM 7[c] [54]	WM 8 [53]	WM 9 [50]	WM 10 [47]	WM 11 [58]
Behavior change theories mentioned used											
Social cognitive theory	—[d]	—	—	—	—	—	—	—	—	—	—
Cognitive behavioral therapy, ABC model	✓[e]	—	—	—	—	—	—	—	—	✓	—
Health action process approach model	✓	—	—	—	—	—	—	—	—	—	—
The transtheoretical model	—	—	—	✓	—	✓	—	—	—	—	—
Goal setting and action theories	—	—	—	—	—	—	✓	—	—	—	—
Self-regulation theory	✓	✓	—	—	—	—	✓	—	—	—	—
Regulatory fit theory	—	—	—	—	—	—	—	—	—	—	—
Control theory	—	—	—	—	—	—	—	—	—	—	—
Self-determination theory	—	—	—	—	—	—	—	—	—	—	—
Social support theories	—	—	—	—	—	✓	—	—	—	—	—
Motivational interviewing	—	—	—	—	—	✓	—	✓	—	—	—
Stroebe's theory on behavior change	—	—	—	—	—	—	—	—	—	—	—
Conservation of resources theory	✓	—	—	—	—	—	—	—	—	—	—
Michie's Behavior Change Wheel framework	—	—	—	—	—	—	—	—	—	—	—
Self-directed behavior change theory	—	—	—	—	—	✓	—	—	—	—	—
Michie's Behavior Change Taxonomy											
Scheduled consequences	—	—	—	—	—	✓	—	—	—	—	—
Reward and threat	—	—	A[f]	—	M[g] / A	—	—	—	—	—	—
Repetition and substitution	✓	✓	—	✓	—	✓	✓	—	✓	✓	✓
Antecedents	✓	✓	—	—	—	—	✓	—	—	✓	—
Associations	A	✓	✓	—	✓	A	A	—	✓	✓	—
Covert learning	—	—	—	—	—	—	—	—	—	—	—
Natural consequences	✓	✓	—	—	—	—	—	—	—	—	—
Feedback and monitoring	M	M	✓	✓	✓	✓	M	M	✓	M	M
Goals and planning	✓	✓	✓	✓	✓	✓	✓	M	✓	✓	✓
Social support	✓	M	✓	✓	—	✓	✓	M / A	✓	—	✓
Comparison of behavior	—	—	—	—	—	—	—	—	—	✓	—
Self-belief	✓	✓	—	—	—	—	—	—	—	—	—
Comparison of outcomes	—	✓	—	—	—	—	✓	—	—	—	—
Identity	—	✓	—	—	—	—	—	—	—	—	—
Shaping knowledge	✓	✓	✓	✓	—	✓	✓	—	✓	✓	✓
Regulations	✓	✓	—	—	—	—	—	—	—	✓	—

[a]Study ID in Multimedia Appendix 5.

[b]WM: weight loss maintenance interventions.

[c]WM7: the intervention [54] was also based on motivational theories, aspects of human motivation, and behavior change, not explicitly described.

[d]No behavior change technique (BCT) or theory was mentioned applied.

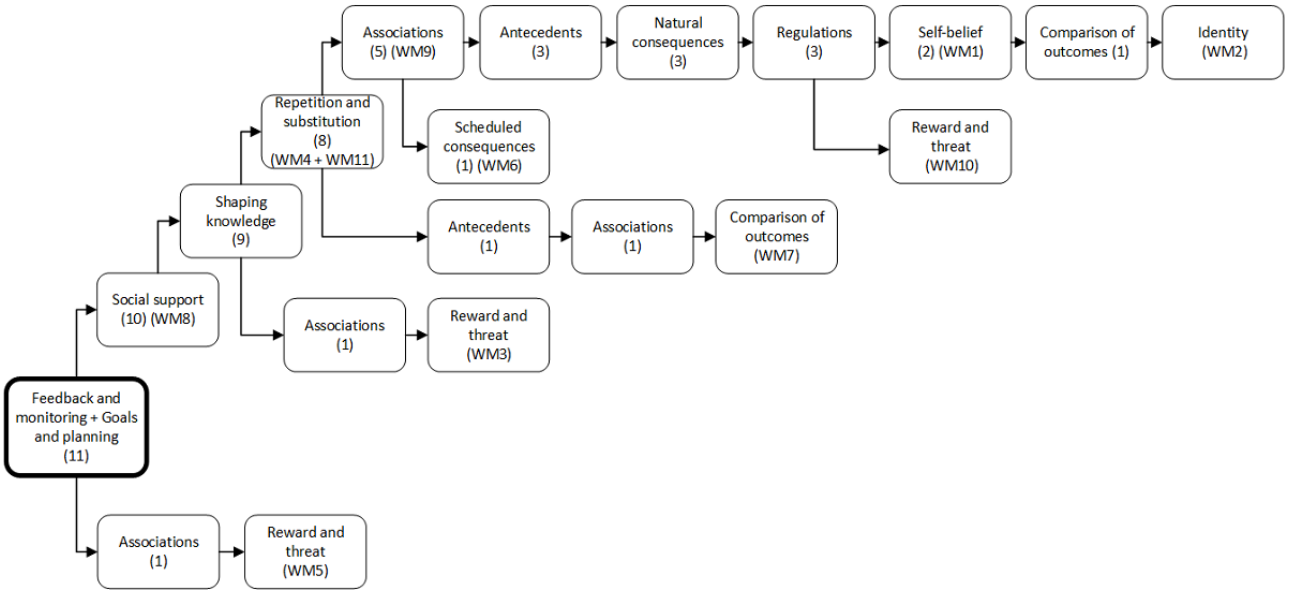
[e]BCT or theory applied.

[f]A: BCT or theory mentioned applied to stimulate adherence.

[g]M: BCT or theory mentioned applied to stimulate motivation.

XSL•FO
RenderX

Figure 3. Flowchart with combinations of behavior change techniques (BCTs) in weight loss maintenance (WM) interventions 1-11. Illustrates the number (n) of WM interventions applying the BCT combined with previous techniques to the left in the flowchart.



Despite linking individual BCTs to stimulation of adherence and/or motivation, as indicated in Table 5, the weight loss maintenance interventions included in the review did not specify the ideal combination of such techniques. The most frequently applied combinations of BCTs in the 11 weight loss maintenance interventions are illustrated in Figure 3. This figure illustrates, from the *bold square* left to right, how many maintenance interventions (n) that actually applied the BCTs, in combination with the previous ones. The combinations *goals and planning* and *feedback and monitoring* were applied in all interventions and were, as indicated in the flowchart (Figure 3), frequently combined with *social support* (91%), *shaping knowledge* (82%), and *repetition and substitution* (73%). A handful of publications described *goals and planning* and *feedback and monitoring* as key strategies or core self-regulation techniques for behavior change and weight outcome and therefore applied these clusters [51,52,54,74,77]. A few publications described *social support* as important for motivation and engagement, reflected in the application of behavioral strategies and the intervention content [51,53]. *Social support*, provided by professionals or peers, was enabled in various ways with or without technology or in combination (blended care) [50,52-54,56,59]. *Shaping knowledge* was present in several ways through offering general suggestions and theory-based advice related to weight loss maintenance [50,51,54,56,58,59]. Finally, the use of *repetition and substitution* typically included habit formation, graded tasks, and behavioral rehearsal [47,51,57,59]. When identifying a specific weight loss maintenance intervention in Figure 3, WM1 to WM11, the exact BCTs applied in the weight loss maintenance intervention can be identified by following the reverse flow to the bold square to the left.

Persuasive System Design Principles Used to Stimulate Motivation and Adherence in Electronic Health Weight Loss Maintenance Interventions

Persuasive System Design Principles and Motivation

The most frequently mentioned PSD principles applied to stimulate motivation in the weight loss maintenance interventions were, as indicated in Table 6, *personalization* (45%), *praise* (45%), and *feedback* (36%). Although *self-monitoring* was used in all included weight loss maintenance interventions and *goal setting* was used in 90% of the interventions, these principles were only mentioned applied once for the purpose to stimulate motivation [53,58].

The weight loss maintenance intervention where the most PSD principles were identified was an advanced gamified smartphone app where characters go through difficult situations, learning to cope with tempting situations (eg, social settings and holidays), and receive *rewards* through healthy habit points [47]. This intervention mentioned stimulating motivation through *tailoring*, *personalization*, and *praise*, for example, motivational messages or cognitive behavioral coping strategies depending on the challenges or situation [47].

Several technologies supporting weight loss maintenance stimulated motivation through motivational *feedback* messages and *praise*, often *personalized* and automated [47,51,53,54,56,57,62]. These messages were applied to motivate the user to stay on course or to provide support on good or bad days, often connected to self-reported feelings, weight (eg, when weight enters a *red zone*), or behaviors related to activities or food. In addition, 1 intervention delivered *tailored*, motivational messages, and coping *suggestions* through gaming elements to learn and *simulate* healthy behaviors [47].

Persuasive System Design Principles and Adherence

Feedback and *rewards* were persuasive principles mentioned to stimulate adherence (eg, 1 intervention used financial *rewards* to stimulate adherence to weekly weight loss maintenance goals) [55]. *Reminders* (27%) were often applied as automated notifications to submit *self-monitoring* information [47,52,57,59] or *remind* users about goals [51,53], although not explicitly mentioned as being applied to stimulate adherence [54,56,57]. *Feedback* and *reminders* (eg, when system usage decreased or when entering of monitoring data was required) and *tailoring* and *personalization* (eg, goal setting and system preferences)

XSL•FO
RenderX

were used to meet the individual needs to stimulate adherence [53,54,56,57].

Persuasive System Design Principles Applied

Several PSD principles were applied in the weight loss maintenance interventions, as presented in Table 6, although usually not explicitly mentioned applied with the purpose of stimulating motivation and adherence in particular.

Table 6. Included weight loss maintenance interventions specifying persuasive system design (PSD) principles.

Study ID	WM[a] 1 [57]	WM 2 [51]	WM 3 [52]	WM 4 [59]	WM 5 [55]	WM 6 [56]	WM 7 [54]	WM 8 [53]	WM 9 [50]	WM 10 [47]	WM 11 [58]
Primary task support											
Reduction	—[b]	—	—	—	—	✓[c]	✓	—	—	✓	—
Tunneling	—	—	—	—	—	✓	✓	—	—	✓	—
Tailoring	✓	✓	✓	✓	✓	✓	✓	M[d] / A[e]	✓	M	✓
Personalization	✓	M	—	—	—	A	M	M	✓	M	✓
Self-monitoring	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	M
Simulation	✓	✓	—	✓	—	✓	✓	✓	✓	✓	—
Rehearsal	—	—	—	—	—	✓	—	—	—	✓	—
Dialog support											
Praise	M	M	—	—	—	—	—	M	—	M	—
Rewards	—	—	A	✓	M / A	✓	✓	—	—	✓	—
Reminders	✓	✓	✓	✓	✓	✓	✓	✓	—	✓	—
Suggestions	✓	—	—	✓	—	✓	✓	✓	—	✓	✓
Similarity	—	—	—	—	—	—	—	—	—	—	—
Liking	—	—	—	—	—	—	—	—	—	—	—
Social role	—	—	—	—	—	—	—	—	—	✓	—
Social support											
Social learning	—	—	—	—	—	—	✓	—	—	—	—
Social comparison	—	—	—	—	—	—	✓	—	—	✓	—
Normative influence	—	—	—	—	—	—	—	—	—	—	—
Social facilitation	—	—	✓	—	—	—	✓	—	—	—	—
Cooperation	—	—	—	—	—	—	✓	—	—	—	—
Competition	—	—	—	—	—	—	—	—	—	—	—
Recognition	—	—	—	—	—	—	—	—	—	✓	—
Other											
Feedback	M / A	M	✓	✓	✓	A	✓	M	✓	M	✓
Goal setting	✓	✓	✓	✓	✓	✓	✓	M	✓	—	✓
Social support	—	✓	✓	—	—	✓	✓	M / A	✓	—	✓

[a]WM: weight loss maintenance intervention.

[b]No PSD was identified applied.

[c]PSD identified.

[d]M: PSD mentioned applied to stimulate motivation.

[e]A: PSD mentioned applied to stimulate adherence.

Operationalization of Commonly Applied Persuasive System Design Principles

Reasons to apply the *self-monitoring* principle were because users could monitor weight, diet, and/or activity, often related to short- or long-term goals [51,53,57], or *self-monitoring* of mood, stress, and/or habit tracking [47,51,53,57]. Monitoring data were registered automatically as well as manually (eg, through wireless trackers and scales) or using the website to enter weight, activity, and diet data [50,54,55]. *Goal setting*

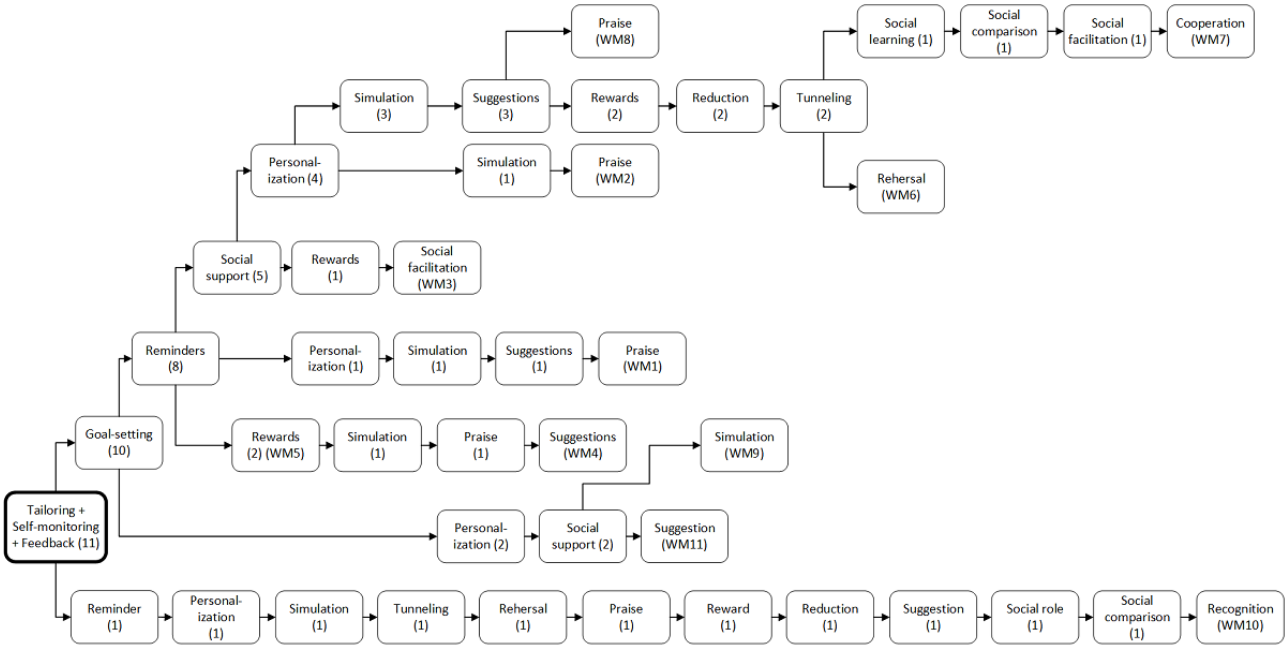
XSL•FO
RenderX

was often integrated into the technology, allowing users to set, monitor, or review both short- and long-term goals (sometimes through *reduction*), related to the behavior they wanted to change [50-56,58,59]. *Tailoring* of messages or *feedback* to the users were again often linked to *self-monitoring* of weight, diet, and activity information, as participants received *tailored feedback* on their progress [51,52,56].

Persuasive System Design Combinations

The most frequent combinations of PSD principles used in all 11 weight loss maintenance interventions were *tailoring*, *self-monitoring*, and *feedback* (100%). As the flowchart in Figure 4 indicates, 90% of the interventions combined these frequently used features with *goal setting* (90%) and *reminders* (73%). The flowchart provides an overview of the number of maintenance interventions (n) that actually applied the various PSD principles, in combination with the previous ones.

Figure 4. Flowchart with combinations of persuasive system design (PSD) principles used in weight loss maintenance (WM) interventions 1-11. Illustrates the number (n) of WM interventions applying the PSD principles combined with previous principles to the left in the flowchart.



Weight Loss Maintenance Effects (Weight Outcomes)

Of the 11 included interventions targeting weight loss maintenance, only 3 evaluated effects of the intervention on body weight or BMI. All 3 interventions found significant effects for weight loss maintenance at 12 weeks postbaseline [59], 12 months postintervention [52], and up to 24 months postbaseline [50].

In addition, one of the interventions had a 6-month weight loss phase followed by a 6-month weight loss maintenance phase and presented characteristics of *high-performing* participants who had lost 10% or greater of their starting weight at the 1-year follow-up [50]. High performers compared with low performers had greater adherence to *self-monitoring* of weight, more days wearing activity trackers, and higher average number of steps per day [50]. In another study, *rewards* were applied to stimulate adherence, suggesting that an internet-delivered cost-benefit approach might be effective to support weight loss maintenance [52]. The third study, entailing a mobile health (ie, mobile or cellular phone technology) lifestyle program, implemented weekly text messages to prevent weight gain, using *praise* to stimulate motivation [59]. This intervention appeared successful in preventing unhealthy weight gain, resulting in modest weight loss and improved health behaviors [59].

The 3 weight loss maintenance interventions achieving effect in terms of weight outcomes were similar in that they all applied techniques and principles related to *tailoring*, *self-monitoring*, *feedback*, *goals and planning*, and *shaping knowledge*. A human *social support* component delivered through a blended format by e-coaching [52], telephone support [59], or expert coach by a Web-based electronic messaging feature [50] was also present in all 3 interventions.

Discussion

Principal Findings

This scoping review aimed to identify BCTs and PSD principles applied in eHealth interventions to support weight loss and weight loss maintenance, as well as techniques and principles applied to stimulate *motivation* and *adherence* for long-term weight loss maintenance. The most successful eHealth weight loss maintenance interventions entailed a combination of BCTs and PSD principles, and the analysis identified several techniques and principles applied to stimulate *motivation* and *adherence*.

Adherence and Motivation

Focus on and description of *motivation* and *adherence* were more prominent in the included weight loss maintenance interventions than in the weight loss interventions. Only 2 of the 45 publications described a definition for *motivation*, and *motivation* was measured in the weight loss interventions but

XSL•FO
RenderX

not in the weight loss maintenance interventions. The results provided some indications that the delivery of tailored, real-time daily feedback messages related to diary entries could enhance motivation [67], use of a virtual coach could be used to motivate users to become more active [49], and autonomous motivation was predictive for adherence to self-monitoring [44]. *Adherence* was defined and measured in various ways in both types of interventions, including but not limited to behavioral adherence, program compliance, technology usage, or adherence to certain technology features (eg, self-monitoring). The evaluation methods applied to measure adherence did not focus on evaluating actual use of the technology but only usage of certain technology features (eg, self-monitoring). The results indicated that *adherence* or usage of self-monitoring techniques was associated with weight loss [50,68,69,74]. The findings related to both *motivation* and *adherence* may provide an interesting input for eHealth development but makes it challenging to compare results across interventions because of diversity in study designs and reporting.

Behavior Change Theories, Behavior Change Techniques, and Persuasive System Design Principles Applied in Weight Loss and Weight Loss Maintenance Interventions

The most frequently used technology in the included interventions of this review was mobile phone, often used for monitoring, dialog, feedback, and support. All interventions had a theoretical anchoring and applied various BCTs and PSD principles. The analysis revealed that techniques and principles applied to support behavior change in weight loss interventions do not necessarily equal weight loss maintenance. However, some key BCTs and PSD principles, identified by applying the Michie's Behavior Change Taxonomy and the PSD model [14,82], including *goal setting*, *self-monitoring*, *feedback*, and *shaping knowledge*, were present in most of the included interventions. The PSD principles from the *primary task support* and *dialog support category* were most frequently applied. *Social support* was also identified as a frequently applied BCT in both types of interventions. Within the PSD model, *social support* was set as a separate PSD principle, as it was difficult to identify within the *social support category* based on the information provided in the publications.

Weight Loss Interventions

The typical weight loss interventions were usually of shorter duration than the weight loss maintenance interventions. *Social cognitive theory* was the most commonly applied behavior change theory in weight loss interventions. The identified core techniques and principles mentioned were used in the technology to support target behavior (weight loss). BCTs and PSD principles more frequently applied in weight loss than weight loss maintenance interventions included *comparison of behavior* and *competition* often to motivate or inspire healthy attitudes and performance between users of the technology (eg, sharing progress, weight changes, and targets achieved).

Weight Loss Maintenance Interventions

The identified 11 weight loss maintenance interventions included in this review had a duration range between 12 and 52 weeks, were often presented as a protocol or described the design and development process only. Of 11 weight loss maintenance interventions, only 3 focused on evaluating weight loss maintenance effects. *Self-regulation theory* was the most often mentioned applied behavior change theory in weight loss maintenance interventions. The core BCTs and PSD principles (eg, *self-monitoring*, *feedback*, *goals and planning*, *tailoring*, and *shaping knowledge*) were reflected in the technology design and considered important for behavior change and weight loss maintenance. Although the ideal combination of BCTs or PSD principles to stimulate motivation, adherence and weight loss maintenance was not explicitly stated, the most frequently mentioned techniques and persuasive principles applied to stimulate motivation were *personalization*, *praise*, and *feedback*, whereas *associations* were frequently mentioned to stimulate adherence. *Rewards* and *social support* were used to stimulate both motivation and adherence. Technologies applying techniques and principles supporting behaviors to deal with biological, environmental, social, behavioral, and cognitive factors (eg, creation of self-determined goals related to healthy habits and self-monitoring) were represented in many of the included weight loss maintenance interventions. In the maintenance phase, s*ocial support*, *rewards*, *reduction*, *praise*, *repetition and substitution*, and *prompts and cues* could be of particular importance in addition to the core techniques identified to address the cost-benefit ratio by incentive driven, rewarding, and persuasive technologies.

The findings in this review are in line with earlier research indicating that behavioral strategies may facilitate health behavior change to maintain weight loss [4,86-88] but that more research focusing on long-term eHealth weight loss maintenance is needed [27,35,89-92].

A recent systematic review on determinants of weight loss maintenance confirmed that evidence related to motivation is sparse [93], and further evidence is needed. Standardization of the adherence concept and reporting [94,95] may also contribute to open the *black box* of eHealth to understand how design and use of eHealth technologies may contribute to improved health and well-being [30].

Lack of information on the *social support* and the *system credibility support* categories have also been pointed to as sparsely reported on by an earlier review on key components in eHealth interventions promoting healthier lifestyle [96]. Earlier research has shown that these categories are important to include when reporting, as users have been less engaged with the technology if credibility was lacking, which again can affect health behavior [97]. This identifies a need for more diligent reporting on design of eHealth interventions and a need for investigation as to which design elements are actually required to achieve behavior change are needed.

This review also shows a lack of user involvement in several of the included interventions. To develop effective eHealth interventions, orchestrated content and system development are needed, as these are often separated by a variety of strategies initiated by researchers and designers of technologies. These challenges can be overcome by multidisciplinary and interwoven

XSL•FO
RenderX

human-centered design approaches during the development of eHealth technologies aimed to change behaviors [12,25].

Although evidence related to theoretical explanation of sustainable maintenance of behavior change is limited [88], existing reviews of technologies point to the need for combinations of BCTs and PSDs to achieve successful health behavior change and weight management [35,91,96,98]. Existing research also points to frequent *self-monitoring* of weight and food intake, high levels of physical activity [87,99,100], and healthy diet as key ingredients often present in weight loss maintenance interventions associated with better weight loss maintenance over time [4,5,8,93,100-103].

Digital developments bring several design opportunities that allow for development and testing of meaningful, adaptive, and sustainable health-promoting solutions [25]. Integration of persuasive interaction and design elements (eg, gaming, avatars, and virtual coach) to reward, rehearse, or simulate cognitive, social, and biological aspects of healthy behaviors or attitudes can provide new methods to learn and maintain new lifestyle and the lost weight, as establishment of healthy behaviors takes time [102]. As smart monitoring is evolving and automatic tracking devices are available in almost all smartphones, this can allow for personalized feedback and long-term monitoring of wellness goals related to a healthy lifestyle that can be maintained lifelong.

Recommendation for Future Design and Research

First, research into design and application of new, personalized digital technologies that integrate sensors and long-term monitoring of data of behaviors and decisions can provide opportunities that may contribute to ultimately solve the conundrum of sustainable health behavior change and long-term weight loss maintenance. Second, the identification of central BCTs and PSD principles to support behavior change, motivation, and adherence in this review allow for user testing in predesign phases of behavioral eHealth interventions, which again can aid in the evaluation of what is needed to truly support individuals in their health. This review also identified self-regulation techniques to support creation and maintenance of healthy habits, but the ideal combination of such techniques should be further investigated through design and evaluation of novel technologies to support long-term weight maintenance after weight loss. Building healthy habits and behaviors takes time, and future research should explore how personalized eHealth technologies can support patients' motivation, long-term adherence, and sustained engagement to improve healthy behaviors over time.

Future research can also better facilitate comparison of interventions through following standardized guidelines and frameworks more diligently when reporting findings, including following guidelines and frameworks such as the CONSORT Guidelines [39], the BCT Taxonomy of Michie [21], and/or the PSD model by Oinas-Kukkonen [16].

Finally, eHealth interventions developed in line with user values and needs may have the potential to motivate and empower sustainable health behavior change, which calls for more user involvement and multidisciplinary approaches in design, development, and evaluation of eHealth interventions. Such an interwoven development process, combining the input, needs, and requirements of researchers, engineers, and stakeholders (including users), is needed to unravel or disentangle the *black box* and create technologies that engage, motivate, and support health behaviors that can be sustained to maintain lost weight for a lifetime.

Strengths and Limitations

This review has a number of limitations. First, identification and scoring of the persuasive features and BCTs could have been prone to subjectivity by the researcher(s). However, to prevent subjectivity, consultation and a 10% validation of included publications were performed. Second, *system credibility support* was not included in the analyses, as limited data were reported and examinations of the actual technology were not included in this review. However, the analysis reveals knowledge about the other 3 PSD categories, in particular, the *primary tasks* and *dialog support*. Third, no quality appraisal of evidence, often done in systematic reviews, was performed in this scoping review. This limits the possibility of drawing conclusions regarding cause and (long-term) effectiveness of interventions. However, as the aim of this scoping review was to provide insight into the design of eHealth interventions, particularly the PSD principles and BCTs mentioned applied to support sustained behavior change and weight loss maintenance, a quality assessment of the studies included was not considered to be as relevant. In addition, insight into various study designs provides an overview over this emerging research area. Finally, the heterogeneity of included designs and lack of long-term results complicate a comparison of the interventions and the possible impact of techniques and principles on reported outcomes, which may introduce bias.

The focus on weight loss maintenance is, however, a major strength of this scoping review, as weight loss maintenance is an area in dire need of further research and future recommendations. In addition, the inclusion of a variety of study designs allows for a consideration of existing interventions that describe design choices and formative evaluations, contributing to eventually opening up the *black box* and giving direction for future design of eHealth interventions.

Conclusions

To the best of our knowledge, this review is the first to identify BCTs and PSD principles applied in eHealth weight loss and weight loss maintenance interventions. Results reveal very limited existing research in the area of eHealth interventions to support weight loss maintenance. *Motivation* and *adherence* are clearly of essence in terms of achieving long-term weight loss maintenance, yet there is still a lack of standardization in definitions and measurement of these concepts. Results show how self-regulation strategies are applied in weight loss and weight loss maintenance interventions, reflected in the design through core techniques and principles such as *self-monitoring*, *feedback*, *goals and planning*, *tailoring*, and *shaping knowledge*. Frequently mentioned BCTs and PSD principles applied to stimulate *motivation* in weight loss maintenance interventions were *personalization*, *praise*, and *feedback*, whereas *associations* were mentioned to stimulate adherence. *Social support* and

XSL•FO
RenderX

rewards were mentioned as being applied to stimulate both motivation and adherence. The most effective combination of techniques or design features to stimulate *motivation*, *adherence*, and weight loss maintenance nevertheless remains somewhat obscure. Although few weight loss maintenance eHealth interventions indicated effect (ie, weight outcome), the interventions with significant results all applied the identified core BCTs and PSD principles, as well as a human *social support* component.

In conclusion, this scoping review aimed to contribute to open the *black box* of eHealth in the design of weight loss maintenance interventions. The findings are expected to contribute to a better understanding of existing research in this field, and in addition to contribute to development and evaluation of future eHealth interventions and novel solutions to support sustained behavior change and long-term weight loss maintenance. The results of this review support the notion that the research of eHealth interventions in weight loss maintenance is still in its infancy, and more research is needed.

Acknowledgments

This study was funded by Vestfold Hospital Trust and Oslo University Hospital in Norway. The authors would like to thank the librarians Mariann Mathisen, MSc, and Julie Skattebu at Vestfold Hospital Trust and Peter Noort, MSc, at University of Twente for their contribution developing the search strategy. They are also very grateful to researcher Floor Sieverink, PhD, for participating in the initial screening process and analysis and valuable consultation during the process by a PhD student Roberto R Cruz Martínez, MSc.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Cluster label and component behavior change techniques (Michie et al, 2013).

[PDF File (Adobe PDF File), 89KB - jmir_v21i6e14265_app1.pdf]

Multimedia Appendix 2

The persuasive system design model (Oinas-Kukkonen, 2009).

[PDF File (Adobe PDF File), 98KB - jmir_v21i6e14265_app2.pdf]

Multimedia Appendix 3

Search strategy.

[PDF File (Adobe PDF File), 73KB - jmir_v21i6e14265_app3.pdf]

Multimedia Appendix 4

Inclusion and exclusion criteria.

[PDF File (Adobe PDF File), 90KB - jmir_v21i6e14265_app4.pdf]

Multimedia Appendix 5

Overview of included publications.

[PDF File (Adobe PDF File), 150KB - jmir_v21i6e14265_app5.pdf]

Multimedia Appendix 6

Reported outcomes.

[PDF File (Adobe PDF File), 130KB - jmir_v21i6e14265_app6.pdf]

Multimedia Appendix 7

Overview over behavior change theories and techniques used in the included eHealth interventions.

[PDF File (Adobe PDF File), 107KB - jmir_v21i6e14265_app7.pdf]

XSL•FO
RenderX

Multimedia Appendix 8

Overview over persuasive system design principles used in the included eHealth interventions.

[PDF File (Adobe PDF File), 117KB - jmir_v21i6e14265_app8.pdf]

References

1. Ng M, Fleming T, Robinson M, Thomson B, Graetz N, Margono C, et al. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet 2014 Aug 30;384(9945):766-781. [doi: 10.1016/S0140-6736(14)60460-8] [Medline: 24880830]
2. GBD 2015 Obesity Collaborators, Afshin A, Forouzanfar MH, Reitsma MB, Sur P, Estep K, et al. Health effects of overweight and obesity in 195 countries over 25 years. N Engl J Med 2017 Dec 6;377(1):13-27 [FREE Full text] [doi: 10.1056/NEJMoa1614362] [Medline: 28604169]
3. Christiansen T, Bruun JM, Madsen EL, Richelsen B. Weight loss maintenance in severely obese adults after an intensive lifestyle intervention: 2- to 4-year follow-up. Obesity (Silver Spring) 2007 Feb;15(2):413-420 [FREE Full text] [doi: 10.1038/oby.2007.530] [Medline: 17299115]
4. Montesi L, El Ghoch M, Brodosi L, Calugi S, Marchesini G, Dalle GR. Long-term weight loss maintenance for obesity: a multidisciplinary approach. Diabetes Metab Syndr Obes 2016;9:37-46 [FREE Full text] [doi: 10.2147/DMSO.S89836] [Medline: 27013897]
5. Garcia UC, Huizinga M, Beech B, Elasy T. American Diabetes Association - Clinical Diabetes Journal.: American Diabetes Association; 2008 Jul 01. Weight Regain Prevention URL: http://clinical.diabetesjournals.org/content/26/3/100
6. MacLean PS, Wing RR, Davidson T, Epstein L, Goodpaster B, Hall KD, et al. NIH working group report: innovative research to improve maintenance of weight loss. Obesity (Silver Spring) 2015 Jan;23(1):7-15 [FREE Full text] [doi: 10.1002/oby.20967] [Medline: 25469998]
7. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
8. American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Expert Panel, 2013. Executive summary: Guidelines (2013) for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Obesity Society published by the Obesity Society and American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Based on a systematic review from the The Obesity Expert Panel, 2013. Obesity (Silver Spring) 2014 Jul;22(Suppl 2):S5-39. [doi: 10.1002/oby.20821] [Medline: 24961825]
9. Lean M, Hankey C. Keeping it off: the challenge of weight-loss maintenance. Lancet Diabetes Endocrinol 2018 Dec;6(9):681-683 [FREE Full text] [doi: 10.1016/S2213-8587(17)30405-9] [Medline: 29371077]
10. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
11. van Gemert-Pijnen JE, Peters O, Ossebaard HC. Improving eHealth. The Hague: Eleven International Publishing; 2013.
12. van Gemert-Pijnen JE, Kelders SM, Kip H, Sanderman R. Ehealth Research, Theory And Development: A Multi-Disciplinary Approach. Abingdon: Routledge; 2018.
13. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
14. Nijland N. Grounding eHealth: Towards a Holistic Framework for Sustainable eHealth Technologies. Enschede: University of Twente; 2011.
15. European Commission. European Commission. 2018. Communication from the Commission to the European Parlament, The Council, The European Economic and Social Committee and The Committee of the Regions on enabling the digital transformation of health and care in the Digital Single Market; empowering citizens and building a healthier society URL: https://tinyurl.com/y7hzukat
16. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model and system features. Commun Assoc Inf Syst 2009;24:28. [doi: 10.17705/1CAIS.02428]
17. 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]
18. Wang J, Cai C, Padhye N, Orlander P, Zare M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. JMIR Mhealth Uhealth 2018 Apr 10;6(4):e92 [FREE Full text] [doi: 10.2196/mhealth.4478] [Medline: 29636320]
19. 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;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]

XSL•FO
RenderX

20. Sorgente A, Pietrabissa G, Manzoni GM, Re F, Simpson S, Perona S, et al. Web-based interventions for weight loss or weight loss maintenance in overweight and obese people: a systematic review of systematic reviews. J Med Internet Res 2017 Jun 26;19(6):e229 [FREE Full text] [doi: 10.2196/jmir.6972] [Medline: 28652225]
21. 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]
22. Kelders SM, Oinas-Kukkonen H, Oörni A, van Gemert-Pijnen JE. Health behavior change support systems as a research discipline; a viewpoint. Int J Med Inform 2016 Dec;96:3-10. [doi: 10.1016/j.ijmedinf.2016.06.022] [Medline: 27432207]
23. Michie S, Abraham C, Eccles MP, Francis JJ, Hardeman W, Johnston M. Strengthening evaluation and implementation by specifying components of behaviour change interventions: a study protocol. Implement Sci 2011;6:10 [FREE Full text] [doi: 10.1186/1748-5908-6-10] [Medline: 21299860]
24. Fogg B. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: 4th International Conference on Persuasive Technology; April 26-29, 2009; Claremont, California, USA p. 1-7 URL: https://dl.acm.org/citation.cfm?id=1541999
25. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 5;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
26. Fogg B. Persuasive Technology: Using Computers To Change What We Think And Do. San Francisco: Morgan Kaufmann; 2002.
27. Beleigoli AM, Andrade AQ, Cançado AG, Paulo MN, Diniz MD, Ribeiro AL. Web-based digital health interventions for weight loss and lifestyle habit changes in overweight and obese adults: systematic review and meta-analysis. J Med Internet Res 2019 Jan 8;21(1):e298 [FREE Full text] [doi: 10.2196/jmir.9609] [Medline: 30622090]
28. Ludden G, Cain R, MacKrill J, Allen F. Design for behaviour change for health and wellbeing. In: Niedderer K, Clune S, Ludden G, editors. Design for Behaviour Change: Theories and Practices of Designing for Change, 1st Edition. Abingdon: Routledge, Taylor and Francis; 2017:184-199.
29. Fjeldsoe B, Neuhaus M, Winkler E, Eakin E. Systematic review of maintenance of behavior change following physical activity and dietary interventions. Health Psychol 2011 Jan;30(1):99-109. [doi: 10.1037/a0021974] [Medline: 21299298]
30. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 7;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
31. Resnicow K, Strecher V, Couper M, Chua H, Little R, Nair V, et al. Methodologic and design issues in patient-centered e-health research. Am J Prev Med 2010 Jan;38(1):98-102. [doi: 10.1016/j.amepre.2009.09.034] [Medline: 20117564]
32. Lehto T, Oinas-Kukkonen H. Persuasive features in six weight loss websites: a qualitative evaluation. In: Persuasive Technology: 5th International Conference, Persuasive 2010, Copenhagen, Denmark, June 7-10, 2010, Proceedings (lecture Notes In Computer Science). Berlin: Springer; 2019.
33. Kelders SM, van Gemert-Pijnen JE. Using log-data as a starting point to make eHealth more persuasive. In: Persuasive Technology: 8th International Conference, Persuasive 2013, Sydney, Nsw, Australia, April 3-5, 2013. Proceedings (lecture Notes In Computer Science). Berlin: Springer; 2013:99-109.
34. Räisänen T, Lehto T, Oinas-Kukkonen H. Practical findings from applying the PSD model for evaluating software design specifications. In: Persuasive Technology: 5th International Conference, Persuasive 2010, Copenhagen, Denmark, June 7-10, 2010, Proceedings (lecture Notes In Computer Science). Berlin: Springer; 2010:185-192.
35. Neve M, Morgan PJ, Jones PR, Collins CE. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. Obes Rev 2010 Apr;11(4):306-321. [doi: 10.1111/j.1467-789X.2009.00646.x] [Medline: 19754633]
36. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
37. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res Synth Methods 2014 Dec;5(4):371-385 [FREE Full text] [doi: 10.1002/jrsm.1123] [Medline: 26052958]
38. Elliot AJ, Covington MV. Approach and avoidance motivation. Edu Psychol Rev 2001:73-92. [doi: 10.1023/A:1009009018235]
39. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
40. 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]

XSL•FO
RenderX

41. Maclean PS, Bergouignan A, Cornier M, Jackman MR. Biology's response to dieting: the impetus for weight regain. Am J Physiol Regul Integr Comp Physiol 2011 Sep;301(3):R581-R600 [FREE Full text] [doi: 10.1152/ajpregu.00755.2010] [Medline: 21677272]
42. Covidence - A Cochrane technology platform. URL: https://www.covidence.org
43. Webber KH, Gabriele JM, Tate DF, Dignan MB. The effect of a motivational intervention on weight loss is moderated by level of baseline controlled motivation. Int J Behav Nutr Phys Act 2010;7:4 [FREE Full text] [doi: 10.1186/1479-5868-7-4] [Medline: 20157441]
44. Webber KH, Tate DF, Ward DS, Bowling JM. Motivation and its relationship to adherence to self-monitoring and weight loss in a 16-week internet behavioral weight loss intervention. J Nutr Educ Behav 2010;42(3):161-167. [doi: 10.1016/j.jneb.2009.03.001] [Medline: 20138583]
45. Lentelink SJ, Spil AA, Broens T, Hermens HJ, Jones VM. Healthy weight game!: Lose weight together: The design and evaluation of a serious game for overweight and obesity. In: 2013 IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH). Piscataway: IEEE; 2013.
46. Matthews L, Pugmire J, Moore L, Kelson M, McConnachie A, McIntosh E, et al. Study protocol for the 'HelpMeDoIt!' randomised controlled feasibility trial: an app, web and social support-based weight loss intervention for adults with obesity. BMJ Open 2017 Oct 25;7(10):e017159 [FREE Full text] [doi: 10.1136/bmjopen-2017-017159] [Medline: 29074513]
47. Podina IR, Fodor LA, Cosmoiu A, Boian R. An evidence-based gamified mHealth intervention for overweight young adults with maladaptive eating habits: study protocol for a randomized controlled trial. Trials 2017 Dec 12;18(1):592 [FREE Full text] [doi: 10.1186/s13063-017-2340-6] [Medline: 29233169]
48. Johnston JD, Massey AP, Devaneaux CA. Innovation in weight loss programs: a 3-dimensional virtual-world approach. J Med Internet Res 2012;14(5):e120 [FREE Full text] [doi: 10.2196/jmir.2254] [Medline: 22995535]
49. Watson A, Bickmore T, Cange A, Kulshreshtha A, Kvedar J. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. J Med Internet Res 2012;14(1):e1 [FREE Full text] [doi: 10.2196/jmir.1629] [Medline: 22281837]
50. Painter S, Ditsch G, Ahmed R, Hanson NB, Kachin K, Berger J. Retrofit weight-loss outcomes at 6, 12, and 24 months and characteristics of 12-month high performers: a retrospective analysis. JMIR Mhealth Uhealth 2016 Aug 22;4(3):e101 [FREE Full text] [doi: 10.2196/mhealth.5873] [Medline: 27549134]
51. Evans EH, Araújo-Soares V, Adamson A, Batterham AM, Brown H, Campbell M, et al. The NULevel trial of a scalable, technology-assisted weight loss maintenance intervention for obese adults after clinically significant weight loss: study protocol for a randomised controlled trial. Trials 2015 Sep 22;16:421 [FREE Full text] [doi: 10.1186/s13063-015-0931-7] [Medline: 26395774]
52. Leahey TM, Fava JL, Seiden A, Fernandes D, Doyle C, Kent K, et al. A randomized controlled trial testing an internet delivered cost-benefit approach to weight loss maintenance. Prev Med 2016 Dec;92:51-57 [FREE Full text] [doi: 10.1016/j.ypmed.2016.04.013] [Medline: 27095323]
53. Barnett J, Harricharan M, Fletcher D, Gilchrist B, Coughlan J. myPace: an integrative health platform for supporting weight loss and maintenance behaviors. IEEE J Biomed Health Inform 2015 Jan;19(1):109-116 [FREE Full text] [doi: 10.1109/JBHI.2014.2366832] [Medline: 25438329]
54. van Genugten L, van Empelen P, Oenema A. Intervention use and action planning in a web-based computer-tailored weight management program for overweight adults: randomized controlled trial. JMIR Res Protoc 2014;3(3):e31 [FREE Full text] [doi: 10.2196/resprot.2599] [Medline: 25057122]
55. Shaw PA, Yancy WS, Wesby L, Ulrich V, Troxel AB, Huffman D, et al. The design and conduct of Keep It Off: an online randomized trial of financial incentives for weight-loss maintenance. Clin Trials 2017 Feb;14(1):29-36 [FREE Full text] [doi: 10.1177/1740774516669679] [Medline: 27646508]
56. Stevens VJ, Funk KL, Brantley PJ, Erlinger TP, Myers VH, Champagne CM, et al. Design and implementation of an interactive website to support long-term maintenance of weight loss. J Med Internet Res 2008 Jan 25;10(1):e1 [FREE Full text] [doi: 10.2196/jmir.931] [Medline: 18244892]
57. Brindal E, Hendrie GA, Freyne J. combining persuasive technology with behavioral theory to support weight maintenance through a mobile phone app: protocol for the MotiMate app. JMIR Res Protoc 2016 Jan 8;5(1):e5 [FREE Full text] [doi: 10.2196/resprot.4664] [Medline: 26747725]
58. Zoppis I, Mauri G, Sicurello F, Santoro E, Castelnuovo G. DIABESITY: Design of mHealth integrated solutions for empowering diabetic and obese citizens in self-monitoring and self-management using mobile devices, apps, social media and web-based technologies. In: de Alencar MS, editor. Communication, Management and Information Technology. Abingdon: Taylor & Francis Group; 2017:207-214.
59. 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]
60. Crane MM, Ward DS, Lutes LD, Bowling JM, Tate DF. Theoretical and behavioral mediators of a weight loss intervention for men. Ann Behav Med 2016 Dec;50(3):460-470 [FREE Full text] [doi: 10.1007/s12160-016-9774-z] [Medline: 26842133]

XSL•FO
RenderX

61. Pellegrini CA, Hoffman SA, Collins LM, Spring B. Optimization of remotely delivered intensive lifestyle treatment for obesity using the Multiphase Optimization Strategy: Opt-IN study protocol. Contemp Clin Trials 2014 Jul;38(2):251-259 [FREE Full text] [doi: 10.1016/j.cct.2014.05.007] [Medline: 24846621]
62. Banos RM, Mensorio MS, Cebolla A, Rodilla E, Palomar G, Lisón J, et al. An internet-based self-administered intervention for promoting healthy habits and weight loss in hypertensive people who are overweight or obese: a randomized controlled trial. BMC Cardiovasc Disord 2015 Aug 4;15:83 [FREE Full text] [doi: 10.1186/s12872-015-0078-1] [Medline: 26239241]
63. Cohen A, Perozich A, Rajan R, Persky S, Parisi J, Bowie J, et al. Framed, interactive theory-driven texting: effects of message framing on health behavior change for weight loss. Fam Community Health 2017;40(1):43-51 [FREE Full text] [doi: 10.1097/FCH.0000000000000128] [Medline: 27870754]
64. Bouhaidar CM, DeShazo JP, Puri P, Gray P, Robins JL, Salyer J. Text messaging as adjunct to community-based weight management program. Comput Inform Nurs 2013 Oct;31(10):469-476. [doi: 10.1097/01.NCN.0000432121.02323.cb] [Medline: 24141325]
65. Mehring M, Haag M, Linde K, Wagenpfeil S, Frensch F, Blome J, et al. Effects of a general practice guided web-based weight reduction program--results of a cluster-randomized controlled trial. BMC Fam Pract 2013;14:76 [FREE Full text] [doi: 10.1186/1471-2296-14-76] [Medline: 23981507]
66. Miller WR, Rollnick S. Motivational Interviewing: Preparing People For Change, 2nd Edition. New York, NY: The Guilford Press; 2002.
67. Ambeba EJ, Ye L, Sereika SM, Styn MA, Acharya SD, Sevick MA, et al. The use of mHealth to deliver tailored messages reduces reported energy and fat intake. J Cardiovasc Nurs 2015;30(1):35-43. [doi: 10.1097/JCN.0000000000000120] [Medline: 24434827]
68. Anton SD, LeBlanc E, Allen HR, Karabetian C, Sacks F, Bray G, et al. Use of a computerized tracking system to monitor and provide feedback on dietary goals for calorie-restricted diets: the POUNDS LOST study. J Diabetes Sci Technol 2012 Sep;6(5):1216-1225 [FREE Full text] [Medline: 23063049]
69. Burke LE, Zheng Y, Ma Q, Mancino J, Loar I, Music E, et al. The SMARTER pilot study: testing feasibility of real-time feedback for dietary self-monitoring. Prev Med Rep 2017 Jun;6:278-285 [FREE Full text] [doi: 10.1016/j.pmedr.2017.03.017] [Medline: 28409090]
70. 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]
71. Héroux M, Watt M, McGuire KA, Berardi JM. A personalized, multi-platform nutrition, exercise, and lifestyle coaching program: a pilot in women. Internet Interv 2017 Mar;7:16-22 [FREE Full text] [doi: 10.1016/j.invent.2016.12.002] [Medline: 30135822]
72. Shapiro JR, Koro T, Doran N, Thompson S, Sallis JF, Calfas K, et al. Text4Diet: a randomized controlled study using text messaging for weight loss behaviors. Prev Med 2012 Nov;55(5):412-417. [doi: 10.1016/j.ypmed.2012.08.011] [Medline: 22944150]
73. Hutchesson MJ, Tan CY, Morgan P, Callister R, Collins C. Enhancement of self-monitoring in a web-based weight loss program by extra individualized feedback and reminders: randomized trial. J Med Internet Res 2016 Apr 12;18(4):e82 [FREE Full text] [doi: 10.2196/jmir.4100] [Medline: 27072817]
74. Krukowski RA, Harvey-Berino J, Bursac Z, Ashikaga T, West DS. Patterns of success: online self-monitoring in a web-based behavioral weight control program. Health Psychol 2013 Feb;32(2):164-170. [doi: 10.1037/a0028135] [Medline: 22545978]
75. Mateo KF, Jay M. Access to a behavioral weight loss website with or without group sessions increased weight loss in statewide campaign. J Clin Outcomes Manage 2014:345-349 [FREE Full text]
76. Mattila E, Pärkkä J, Hermersdorf M, Kaasinen J, Vainio J, Samposalo K, et al. Mobile diary for wellness management--results on usage and usability in two user studies. IEEE Trans Inf Technol Biomed 2008 Jul;12(4):501-512. [doi: 10.1109/TITB.2007.908237] [Medline: 18632330]
77. Pagoto S, Tulu B, Agu E, Waring ME, Oleski JL, Jake-Schoffman DE. Using the Habit App for weight loss problem solving: development and feasibility study. JMIR Mhealth Uhealth 2018 Jun 20;6(6):e145 [FREE Full text] [doi: 10.2196/mhealth.9801] [Medline: 29925496]
78. Ma J, Xiao L, Blonstein AC. Measurement of self-monitoring web technology acceptance and use in an e-health weight-loss trial. Telemed J E Health 2013 Oct;19(10):739-745 [FREE Full text] [doi: 10.1089/tmj.2013.0009] [Medline: 23952787]
79. Gabriele JM, Carpenter BD, Tate DF, Fisher EB. Directive and nondirective e-coach support for weight loss in overweight adults. Ann Behav Med 2011 Apr;41(2):252-263 [FREE Full text] [doi: 10.1007/s12160-010-9240-2] [Medline: 21108032]
80. Merchant G, Weibel N, Pina L, Griswold WG, Fowler JH, Ayala GX, et al. Face-to-face and online networks: college students' experiences in a weight-loss trial. J Health Commun 2017 Dec;22(1):75-83 [FREE Full text] [doi: 10.1080/10810730.2016.1250847] [Medline: 28060581]
81. Rutledge T, Skoyen JA, Wiese JA, Ober KM, Woods GN. A comparison of MOVE! versus TeleMOVE programs for weight loss in Veterans with obesity. Obes Res Clin Pract 2017;11(3):344-351. [doi: 10.1016/j.orcp.2016.11.005] [Medline: 27931766]

XSL•FO
RenderX

82. Sevick MA, Woolf K, Mattoo A, Katz SD, Li H, St-Jules DE, et al. The Healthy Hearts and Kidneys (HHK) study: Design of a 2×2 RCT of technology-supported self-monitoring and social cognitive theory-based counseling to engage overweight people with diabetes and chronic kidney disease in multiple lifestyle changes. Contemp Clin Trials 2018 Dec;64:265-273 [FREE Full text] [doi: 10.1016/j.cct.2017.08.020] [Medline: 28867396]
83. Alnasser A, Sathiaseelan A, Al-Khalifa A, Marais D. Development of 'Twazon': an Arabic app for weight loss. JMIR Res Protoc 2016 May 16;5(2):e76 [FREE Full text] [doi: 10.2196/resprot.5497] [Medline: 27185568]
84. Leahey TM, Kumar R, Weinberg BM, Wing RR. Teammates and social influence affect weight loss outcomes in a team-based weight loss competition. Obesity (Silver Spring) 2012 Jul;20(7):1413-1418 [FREE Full text] [doi: 10.1038/oby.2012.18] [Medline: 22310234]
85. Limam Mansar S, Jariwala S, Behih N, Shahzad M, Anggraini A. Adapting a database of text messages to a mobile-based weight loss program: the case of the middle East. Int J Telemed Appl 2014;2014:658149 [FREE Full text] [doi: 10.1155/2014/658149] [Medline: 24511311]
86. Gilmartin J, Murphy M. The effects of contemporary behavioural weight loss maintenance interventions for long term weight loss: a systematic review. J Res Nurs 2015;20(6):481-496 [FREE Full text] [doi: 10.1177/1744987115599671]
87. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Sniehotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. Br Med J 2014;348:g2646 [FREE Full text] [Medline: 25134100]
88. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. Health Psychol Rev 2016 Sep;10(3):277-296 [FREE Full text] [doi: 10.1080/17437199.2016.1151372] [Medline: 26854092]
89. 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]
90. Wang Y, Xue H, Huang Y, Huang L, Zhang D. A systematic review of application and effectiveness of mHealth interventions for obesity and diabetes treatment and self-management. Adv Nutr 2017 May;8(3):449-462. [doi: 10.3945/an.116.014100] [Medline: 28507010]
91. Coons MJ, Demott A, Buscemi J, Duncan JM, Pellegrini CA, Steglitz J, et al. Technology interventions to curb obesity: a systematic review of the current literature. Curr Cardiovasc Risk Rep 2012 Apr;6(2):120-134 [FREE Full text] [doi: 10.1007/s12170-012-0222-8] [Medline: 23082235]
92. Manzoni GM, Pagnini F, Corti S, Molinari E, Castelnuovo G. Internet-based behavioral interventions for obesity: an updated systematic review. Clin Pract Epidemiol Ment Health 2011 Mar;7:19-28 [FREE Full text] [doi: 10.2174/1745017901107010019] [Medline: 21552423]
93. Varkevisser RDM, van Stralen MM, Kroeze W, Ket JC, Steenhuis IH. Determinants of weight loss maintenance: a systematic review. Obes Rev 2019 Dec;20(2):171-211 [FREE Full text] [doi: 10.1111/obr.12772] [Medline: 30324651]
94. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 6;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
95. Payne JE, Turk MT, Kalarchian MA, Pellegrini CA. Defining adherence to dietary self-monitoring using a mobile app: a narrative review. J Acad Nutr Diet 2018 Nov;118(11):2094-2119 [FREE Full text] [doi: 10.1016/j.jand.2018.05.011] [Medline: 30115555]
96. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive eCoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017 Aug 1;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
97. Harris PR, Sillence E, Briggs P. The effect of credibility-related design cues on responses to a web-based message about the breast cancer risks from alcohol: randomized controlled trial. J Med Internet Res 2009 Aug 25;11(3):e37 [FREE Full text] [doi: 10.2196/jmir.1097] [Medline: 19709989]
98. Carter DD, Robinson K, Forbes J, Hayes S. Experiences of mobile health in promoting physical activity: a qualitative systematic review and meta-ethnography. PLoS One 2018;13(12):e0208759 [FREE Full text] [doi: 10.1371/journal.pone.0208759] [Medline: 30557396]
99. Jakicic JM, Winters C, Lang W, Wing RR. Effects of intermittent exercise and use of home exercise equipment on adherence, weight loss, and fitness in overweight women: a randomized trial. J Am Med Assoc 1999 Oct 27;282(16):1554-1560. [Medline: 10546695]
100. Jeffery RW, Wing RR, Sherwood NE, Tate DF. Physical activity and weight loss: does prescribing higher physical activity goals improve outcome? Am J Clin Nutr 2003 Oct;78(4):684-689 [FREE Full text] [doi: 10.1093/ajcn/78.4.684] [Medline: 14522725]
101. Thomas JG, Bond DS, Phelan S, Hill JO, Wing RR. Weight-loss maintenance for 10 years in the National Weight Control Registry. Am J Prev Med 2014 Jan;46(1):17-23. [doi: 10.1016/j.amepre.2013.08.019] [Medline: 24355667]
102. Wing RR, Phelan S. Long-term weight loss maintenance. Am J Clin Nutr 2005 Jul;82(1 Suppl):222S-225S [FREE Full text] [Medline: 16002825]

XSL•FO
RenderX

103. Elfhag K, Rössner S. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. Obes Rev 2005 Feb;6(1):67-85. [doi: 10.1111/j.1467-789X.2005.00170.x] [Medline: 15655039]

Abbreviations

BCT: behavior change techniques
BMI: body mass index
eHealth: electronic health
PSD: persuasive system design

Edited by G Eysenbach; submitted 06.04.19; peer-reviewed by T Solvoll, S Langrial, A Beleigoli; comments to author 27.04.19; revised version received 10.05.19; accepted 11.05.19; published 21.06.19.

Please cite as:
Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmesæth J, van Gemert-Pijnen JEWC
Persuasive System Design Principles and Behavior Change Techniques to Stimulate Motivation and Adherence in Electronic Health Interventions to Support Weight Loss Maintenance: Scoping Review
J Med Internet Res 2019;21(6):e14265
URL: http://www.jmir.org/2019/6/e14265/
doi:10.2196/14265
PMID:31228174

©Rikke Aune Asbjørnsen, Mirjam Lien Smedsrød, Lise Solberg Nes, Jobke Wentzel, Cecilie Varsi, Jøran Hjelmesæth, Julia EWC van Gemert-Pijnen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association Between Health Literacy, Electronic Health Literacy, Disease-Specific Knowledge, and Health-Related Quality of Life Among Adults With Chronic Obstructive Pulmonary Disease: Cross-Sectional Study

Michael Stellefson[1], PhD, MCHES; Samantha R Paige[2], PhD, MPH; Julia M Alber[3], PhD, MPH; Beth H Chaney[1], PhD, MCHES; Don Chaney[1], PhD, MCHES; Avery Apperson[1], BS; Arjun Mohan[4], MD

[1]Department of Health Education and Promotion, East Carolina University, Greenville, NC, United States

[2]STEM Translational Communication Center, College of Journalism and Communications, University of Florida, Gainesville, FL, United States

[3]Department of Kinesiology and Public Health, College of Science & Mathematics, California Polytechnic State University, San Luis Obispo, CA, United States

[4]Division of Pulmonary, Critical Care and Sleep Medicine, Department of Internal Medicine, East Carolina University Brody School of Medicine, Greenville, NC, United States

Corresponding Author:
Michael Stellefson, PhD, MCHES
Department of Health Education and Promotion
East Carolina University
3202 Carol G Belk Building
Greenville, NC, 27858
United States
Phone: 1 252 328 2105
Fax: 1 252 328 1285
Email: stellefsonm17@ecu.edu

Abstract

Background: Despite the relatively high prevalence of low health literacy among individuals living with chronic obstructive pulmonary disease (COPD), limited empirical attention has been paid to the cognitive and health literacy–related skills that can uniquely influence patients' health-related quality of life (HRQoL) outcomes.

Objective: The aim of this study was to examine how health literacy, electronic health (eHealth) literacy, and COPD knowledge are associated with both generic and lung-specific HRQoL in people living with COPD.

Methods: Adults from the COPD Foundation's National Research Registry (n=174) completed a cross-sectional Web-based survey that assessed sociodemographic characteristics, comorbidity status, COPD knowledge, health literacy, eHealth literacy, and generic/lung-specific HRQoL. Hierarchical linear regression models were tested to examine the roles of health literacy and eHealth literacy on generic (model 1) and lung-specific (model 2) HRQoL, after accounting for socioeconomic and comorbidity covariates. Spearman rank correlations examined associations between ordinal HRQoL items and statistically significant hierarchical predictor variables.

Results: After adjusting for confounding factors, health literacy, eHealth literacy, and COPD knowledge accounted for an additional 9% of variance in generic HRQoL (total adjusted R^2=21%; $F_{9,164}$=6.09, P<.001). Health literacy (b=.08, SE 0.02, 95% CI 0.04-0.12) was the only predictor positively associated with generic HRQoL (P<.001). Adding health literacy, eHealth literacy, and COPD knowledge as predictors explained an additional 7.40% of variance in lung-specific HRQoL (total adjusted R^2=26.4%; $F_{8,161}$=8.59, P<.001). Following adjustment for covariates, both health literacy (b=2.63, SE 0.84, 95% CI 0.96-4.29, P<.001) and eHealth literacy (b=1.41, SE 0.67, 95% CI 0.09-2.73, P<.001) were positively associated with lung-specific HRQoL. Health literacy was positively associated with most lung-specific HRQoL indicators (ie, cough frequency, chest tightness, activity limitation at home, confidence leaving home, sleep quality, and energy level), whereas eHealth literacy was positively associated with 5 of 8 (60%) lung-specific HRQoL indicators. Upon controlling for confounders, COPD knowledge (b=−.56, SE 0.29, 95% CI −1.22 to −0.004, P<.05) was inversely associated with lung-specific HRQoL.

XSL•FO
RenderX

Conclusions: Health literacy, but not eHealth literacy, was positively associated with generic HRQoL. However, both health literacy and eHealth literacy were positively associated with lung-specific HRQoL, with higher COPD knowledge indicative of lower lung-specific HRQoL. These results confirm the importance of considering health and eHealth literacy levels when designing patient education programs for people living with COPD. Future research should explore the impact of delivering interventions aimed at improving eHealth and health literacy among patients with COPD, particularly when disease self-management goals are to enhance HRQoL.

(J Med Internet Res 2019;21(6):e12165) doi:10.2196/12165

KEYWORDS

COPD; eHealth; health-related quality of life; health literacy; patient education; health status; internet

Introduction

Background

Chronic obstructive pulmonary disease (COPD) is a leading cause of death in the United States, affecting up to 15 million US adults [1,2]. An estimated 65 million people across the globe are living with moderate to very severe COPD [3] contributing to significant morbidity and mortality [4]. The most significant symptoms in patients with severe COPD include shortness of breath (dyspnea), anxiety, depression, and sleep disturbances [5]. Nearly 1 million patients with COPD are hospitalized annually for acute disease exacerbations, with 20% of hospitalized patients readmitted within 30 days of being discharged [6-8]. Readmission is often because of low levels of hospital discharge readiness and confusion regarding discharge instructions [9]. Compared with other chronic diseases such as heart disease and diabetes, treatment adherence in COPD is lower [10], with only about half of medication doses taken as prescribed to treat dyspnea [11]. Many patients with COPD as well as their families have difficulty understanding medication directions and self-management guidelines, which inhibits them from experiencing the health benefits of current medical treatments. People living with COPD report low knowledge in several self-management domains [12], which leads to uncertainty about how to live with their condition [13]. Consequently, patients with COPD describe feeling frustrated, trapped, and socially isolated by their disease [14].

Living with COPD can significantly affect physical, mental, and emotional domains of health-related quality of life (HRQoL) [15]; however, the impact of COPD on HRQoL varies across age categories and presence of comorbidities. For example, among patients living with COPD, age tends to be inversely related to HRQoL, with younger patients faring worse on various HRQoL domains [15]. Similar levels of dyspnea produce greater influence on HRQoL among middle-aged adults as compared with older adults living with COPD. People living with COPD also experience significant burden as a result of comorbidities [16,17], which appear to contribute to poor clinical outcomes [18,16]. A history of previous breathing exacerbations represents another major determinant of lung-specific HRQoL among patients with stable COPD [16]. Surprisingly, only about one-third of patients create an action plan with their provider(s) on how to address potential breathing exacerbations [12].

Patients living with COPD generally receive little guidance in terms of how to recognize and avoid breathing exacerbations. Most patients can only answer about two-thirds of questions about their condition correctly [19], and 16% do not even know what an exacerbation is [20]. Inadequate access to disease-specific education on self-management is a major problem in COPD [21], which hinders the patients' ability to manage symptoms and utilize health care effectively. Areas of knowledge deficiency include inadequate understanding of lung functioning and lack of information on lifestyle factors affected by COPD. Patients who are younger, attend pulmonary rehabilitation, and have more time since their COPD diagnosis have higher knowledge than older patients without access to pulmonary rehabilitation [21]. Nevertheless, patients who report attendance at pulmonary rehabilitation are only able to answer a few more knowledge questions correctly, on average, as compared with patients not enrolled in pulmonary rehabilitation [21].

Low health literacy, or reduced capacity to understand, evaluate, and act on health information, is associated with worse self-management, more severe COPD, learned helplessness, and lower respiratory-specific HRQoL [22,23]. Puente-Maestue and colleagues [24] found that over 50% of COPD patients have low health literacy, which is often compounded by cognitive impairments resulting from hypoxemia as well as by secondary comorbidities, such as depression and anxiety, which influence learning, comprehension, and decision making [23]. COPD patients with inadequate health literacy visit the emergency room and hospital more often, report greater difficulties carrying out their daily living activities, are more dependent on others, and experience higher rates of multimorbidity [24,25]. Low HRQoL places significant health and economic burdens on these patients as well as the health care system. Furthermore, prior research indicates low health literacy levels and lack of disease-specific knowledge also contribute to elevated emergency room and hospital visits [22], whereas higher levels of eHealth literacy result in greater patient knowledge of their diagnosis and better self-management behaviors [26]. In addition, regardless of lung functioning, patients with low health literacy scored worse on patient-reported outcomes. Limited health literacy can also compromise patient-provider communication about medication management, which interferes with a patient's ability to acquire and use prescribed medications [24,25,27,28].

Although there are multiple communication modes and channels for COPD patient education, many patients are left to primarily locate answers to their disease-related questions outside of primary practice settings through self-directed sources such as the internet and Web-based support groups [26,29]. A growing

XSL•FO
RenderX

number of patients are beginning to access social media and mobile health apps for disease management support [30]. For example, 1 study found that almost 80% of patients living with COPD maintain an active social media account [20]. Accordingly, pharmaceutical companies, medical equipment companies, and hospital systems now sponsor numerous online communities and social media websites dedicated to COPD management. Popular social media websites, such as YouTube, demonstrate some potential to positively reach and engage patients living with COPD [31]; however, educational content and quality on social media varies [32]. Moreover, despite the increased availability of Web-based health information on COPD, patients still experience difficulties accessing evidence-based health information about their condition on the internet. Many patients diagnosed with COPD lack confidence in their ability to distinguish between high- and low-quality sources of Web-based health information [26]. Measuring eHealth literacy, or the ability to find, understand, evaluate, and apply health information from the internet to address health problems, is becoming increasingly important as patients with COPD continue to seek health education and medical advice from various internet-based sources.

Purpose

Although lower health literacy is associated with poorer HRQoL in patients with COPD, we do not yet know which aspects of HRQoL may be most impacted by inadequate health literacy skills. Despite the relatively high prevalence of low health literacy among individuals with COPD, limited empirical attention has been paid to the cognitive and health literacy skills that can uniquely influence patients' general health status (generic HRQoL) and health outcomes faced primarily by people living with COPD (lung-specific HRQoL). Therefore, it remains unclear whether distinct types of health literacy are associated with both generic and lung-specific HRQoL among adults living with COPD. In addition, very few studies have examined the role that eHealth literacy may play in affecting HRQoL in people living with COPD [26]. The potential moderating role of COPD knowledge on relationships between health/eHealth literacy and HRQoL is also yet to be explored. It is expected that higher COPD knowledge could significantly increase (or enhance) the effects of health literacy and eHealth literacy on lung-specific HRQoL. The purpose of this study was to assess how health literacy, eHealth literacy, and COPD-related knowledge are associated with both generic and lung-specific HRQoL in people living with COPD, particularly after accounting for factors known to be associated with HRQoL, such as socioeconomic status and living with a comorbidity. The following 4 hypotheses were tested during this study:

1. Hypothesis 1: Health literacy, eHealth literacy, and COPD knowledge will be positively associated with generic HRQoL among patients living with COPD, even after controlling for potentially confounding factors affecting HRQoL.
2. Hypothesis 2: Health literacy, eHealth literacy, and COPD knowledge will be positively associated with lung-specific HRQoL among patients living with COPD, even after controlling for potentially confounding factors affecting HRQoL.
3. Hypothesis 3: COPD knowledge will moderate the effect of health literacy on lung-specific HRQoL, such that higher COPD knowledge will significantly increase (or enhance) the effect of health literacy on lung-specific HRQoL.
4. Hypothesis 4: COPD knowledge will moderate the effect of eHealth literacy on lung-specific HRQoL, such that higher COPD knowledge will significantly increase (or enhance) the effect of health literacy on lung-specific HRQoL.

Methods

Sample

Adults from the COPD Foundation's National Research Registry were recruited to participate. This confidential database includes individuals from all 50 states of the United States who self-report physician or specialist-diagnosed COPD. Enrolled individuals can opt-in to being contacted about potential research opportunities that are promoted for patients enrolled in the Registry. Patients who choose to opt-in complete informed consent documentation indicating that they agree to be contacted about potential patient-oriented research opportunities. Enrolled adults older than 40 years with a valid email address were eligible to participate in this study (N=1270). COPD develops most often in adults older than 40 years [33]. Over three-quarters of eligible patients were older than 60 years (981/1270, 77.2%), and over half were female (703/1270, 55.35%). Most eligible patients were white (1184/1270, 93.23%). Patients who self-reported being unable to speak English were excluded because of lack of access to linguistically-appropriate instrumentation for all measures.

Procedures

Eligible participants received an email invitation to participate in a Web-based survey powered by Qualtrics. All eligible participants were required to provide electronic informed consent before completing the survey. Institutional Review Board (IRB) approval for the study was obtained from both a hospital-based regulatory body overseeing studies of Registry members, as well as a university-sponsored IRB. Participants were notified that their email address would be entered a random drawing to receive a small electronic gift card for their participation. Due to restrictions stipulating that registrants only be contacted once about opportunities to participate in research, the investigators were unable to recontact nonresponders (n=1071).

Measures

Demographic questions assessed several sociodemographic characteristics, including age (years), sex, race (white, black or African-American, Asian or Pacific Islander, American Indian or Alaska Native, Mixed Race, or other), marital status (now married, now widowed, never married, divorced, or separated), and education level (less than 8th grade, 8th–11th grade, completed high school, some college, college graduate, or postgraduate). Participants were also asked to affirm (yes/no) if they were currently living with 1 or more comorbidities (diabetes, hypertension, heart disease, cancer, stroke-related symptoms, arthritis or joint problems, or other health problems

XSL•FO
RenderX

or conditions). Respondents reporting any of these diseases/disorders were coded as *living with comorbidities*, whereas respondents indicating none of these other conditions were marked as *not living with comorbidities*.

COPD knowledge was measured using the Chronic Obstructive Pulmonary Disease Knowledge Questionnaire (COPD-Q) [34]. The COPD-Q is a valid and reliable 13-item self-administered true/false questionnaire assessing COPD knowledge in patients, regardless of their health literacy. This instrument aids researchers and clinicians in identifying gaps in patients' knowledge about COPD. COPD-Q scores demonstrate acceptable internal consistency (Cronbach alpha=.72) and are significantly related to health literacy and educational attainment [35]. The Spanish version of the COPD-Q has also demonstrated good internal consistency (Cronbach alpha=.85) and high reliability, with an intraclass correlation efficient of .84 for the total score [36].

The 3-item Health Literacy Screening Questionnaire (HLSQ) [37] was used to measure general health literacy. Use of screening questions to assess general health literacy is effective in detecting inadequate health literacy; these items have also been validated against widely used measures of health literacy [38-40]. HLSQ items assess the degree to which participants possess (1) problems with learning health information, (2) self-efficacy completing health-related forms, and (3) problems reading health information on Likert scales ranging from 1 (problems with health literacy) to 5 (no problems with health literacy). HLSQ scores in this study demonstrated marginal internal consistency (Cronbach alpha=.64).

eHealth literacy was assessed using an 8-item rating scale called the eHealth Literacy Scale (eHEALS) [41]. This brief scale assesses one's perceived ability to find, understand, and appraise health information from Web-based sources and apply knowledge to address health concerns. eHEALS is a reliable computer-based measure of patients' knowledge and self-efficacy for obtaining and evaluating Web-based health resources [42]. Responders are asked to indicate their level of agreement with 8 statements describing their Web-based health information–seeking experiences on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Mean Likert scale responses were computed for the 8-item eHEALS. Data collected using the eHEALS in this study showed high internal consistency (Cronbach alpha=.90).

Generic HRQoL was measured using the EuroQol (EQ)-5D [43], an instrument that assesses the following 5 relevant domains of HRQoL: (1) mobility, (2) self-care, (3) usual activity, (4) pain, and (5) anxiety/depression. In each domain, participants are asked to select from 3 responses ranging from *no problems* (level 1) to *some problems* (level 2) to *inability or extreme difficulty* (level 3). Responses are used to quantify an individual's unique health state by using a 5-digit descriptor ranging from 11111 for perfect health to 33333 for worst possible health. Weights are applied to score responses on each of the 5 domains and then converted to a single index value ranging from 0 to 1, where a score of 1 represents a perfect health state. This index value is based on a value set derived from a Time Trade Off valuation study reflecting data collected in the US population [44]. For post hoc correlational analyses, item scores on each EuroQol-5D (EQ-5D) dimension were reverse scored to facilitate result interpretability (ie, *no problems* originally marked as level 1 were recoded to *3* to reflect higher self-report on HRQoL dimension). The EQ-5D is self-administered, written at a 7th-grade level and can be completed in 5 to 10 min [44]. Data collected with the EQ-5D in this study showed satisfactory internal consistency (Cronbach alpha=.74).

Lung-specific HRQoL was assessed using the COPD Assessment Test (CAT) [45]. CAT is a widely used 8-item self-reported questionnaire that quantifies the impact of COPD symptoms. Respondents are asked to self-report the extent to which they are affected by coughing, chest congestion, chest tightness, breathlessness, difficulty leaving their home, sleep quality, and low energy levels. Item responses range from 1 (no problems) to 5 (significant difficulties). For result interpretability, responses to all 8 items were reverse scored before computing the total CAT scores for each patient. After reverse scoring CAT items, lower scores indicated greater difficulties on relevant dimensions of lung-specific HRQoL (ie, score of *1* indicated significant difficulties on dimension of lung-specific HRQoL). Total possible CAT scores can range from 0 (mild) to 40 (very severe) [45,46]. Among clinically stable COPD patients, total CAT scores are highly correlated with scores on the St. George's Respiratory Questionnaire [45], which is another commonly used, yet lengthy (50-item), instrument that measures lung-specific HRQoL in people with obstructive airway disease.

Data Analysis

Frequency statistics were used to summarize categorical data, whereas means (SDs) were computed to describe interval data. Missing items from multi-item scales were mean-imputed if at least half of the scale items were completed. Correlations between main study variables were analyzed using Pearson *r* correlation coefficients.

Hierarchical linear regressions were computed to examine the roles of health literacy and eHealth literacy on generic (model 1) and lung-specific (model 2) HRQoL by sequentially adding predictors into 2 blocks within each model. To control for the effects of demographic covariates and comorbidity status on the dependent variables of generic and lung-specific HRQoL, sociodemographic factors (ie, age, gender, race, ethnicity, marital status, and education level) and comorbidity (present/not present) were entered in block 1 as potential confounding factors affecting each type of HRQoL. Because we were primarily interested in the additional effects of COPD-related knowledge, health literacy, and eHealth literacy above and beyond these covariates, predictor variables were subsequently entered in block 2 of each model. To determine whether health literacy, eHealth literacy, and COPD knowledge provided any significant increment in the amount of variance explained in generic (model 1) and lung-specific HRQoL (model 2), *F*-test statistics were evaluated to determine statistically significant R^2 changes in explained variance (%) at each step of the analysis. Unstandardized beta coefficients were examined for each control/predictor variable, and interaction effects were assessed

XSL•FO
RenderX

by interpreting multiple R^2 and adjusted R^2. Tolerance and variance inflation factor (VIF) estimates for each hierarchical model were computed to ensure that tolerance estimates were below 0.10 and VIF were less than 10 [47]. In addition, Cohen f^2 is a useful measure of local effect size appropriate for hierarchical regression models [48], and this was calculated for each hierarchical linear regression in this study. According to guidelines by Cohen [49], $f^2 \geq 0.02$, $f^2 \geq 0.15$, and $f^2 \geq 0.35$ represent small, medium, and large effect sizes, respectively.

Associations between ordinally scaled HRQoL items and statistically significant predictor variables from hierarchical analyses were assessed using Spearman rank correlation procedures. In the third (and final) step of the hierarchical regression analyses, interaction terms of eHealth literacy×COPD knowledge and health literacy*COPD knowledge were entered in both models, one predicting generic HRQoL and another predicting lung-specific HRQoL. All statistical analyses were conducted using IBM SPSS Statistics version 24 (SPSS Inc, Chicago, IL, USA), with values of $P<.05$ considered statistically significant.

Results

Respondent Characteristics

Overall, 199 adults with COPD submitted responses to the Web-based survey. List-wise deletion was used to remove cases that contained missing data (n=25); therefore, responses from 174 of 1270 patients (response rate=13.70%) were analyzed. Table 1 presents overall characteristics of the study sample (n=174). Participants reported a mean age of 66.06 years (SD 9.43). Slightly over half of the sample were female (88/174, 50.6%), and the clear majority were white (168/174, 96.6%). The sample was also well-educated, with over 80% (147/174, 84.5%) of the sample reporting at least some college education. Most participants were married (106/174, 60.9%); however, almost one-quarter were divorced or separated (42/174, 24.1%).

Table 1. Overall characteristics of the study sample (N=174).

Variable	Study participants
Age (years), mean (SD)	66.06 (9.43)
Sex, n (%)	
Female	88 (50.6)
Male	86 (49.4)
Race, n (%)	
White	168 (96.6)
Black	3 (1.7)
Asian or Pacific Islander	1 (0.6)
American Indian or Alaska Native	1 (0.6)
Mixed race	1 (0.6)
Ethnicity, n (%)	
Hispanic	4 (2.3)
Non-Hispanic	170 (97.7)
Education level, n (%)	
8th to 11th grade	5 (2.9)
High school graduate	22 (12.6)
Some college	83 (47.7)
College graduate	37 (21.3)
Postgraduate degree	27 (15.5)
Marital status, n (%)	
Married	106 (60.9)
Widowed	16 (9.2)
Never married	10 (5.7)
Divorced or separated	42 (24.1)
Comorbidity, n (%)	
Yes	143 (82.2)
No	31 (17.8)

XSL•FO
RenderX

Health Literacy, Electronic Health Literacy, and Chronic Obstructive Pulmonary Disease Knowledge

The mean score on HLSQ items measuring health literacy was 4.52 (SD 0.62), which was higher than self-reported scores on the 8-item eHEALS measure of eHealth literacy (mean 3.63, SD 0.71). On average, respondents answered almost 10 (mean 9.81, SD 1.63) out of 13 questions correctly on the COPD-Q measuring knowledge related to COPD.

Health Status Measures

The mean EQ-5D score was 0.723 (SD 0.18) on the EQ-5D index scale measuring generic HRQoL from 0 to 1, whereas the mean score on the CAT measuring lung-specific HRQoL on a scale of 0 to 40 was 24.02 (SD 6.65), indicating moderate lung-specific HRQoL. EQ-5D index scores were positively associated with CAT scores (r=.61, P<.001). More than 80% of the sample (143/174, 82.2%) reported living with at least one comorbidity.

Hypothesis Testing

Test of Hypothesis 1: Health Literacy, Electronic Health Literacy, and Chronic Obstructive Pulmonary Disease Knowledge on Generic Health-Related Quality of Life

In the hierarchical regression model examining effects of health literacy, eHealth literacy, and COPD knowledge on generic HRQoL, demographic characteristics and comorbidity (step 1) accounted for 16.0% (adjusted R^2=13%) of the variance in generic HRQoL, $F_{6,167}$=5.30, P<.001. Table 2 shows that older age (b=.01, SE 0.001, P<.01) and higher education level (b=.04, SE 0.13, P<.01) were associated with better generic HRQoL. Introducing health literacy, eHealth literacy, and COPD knowledge as predictor variables (step 2) explained an additional 9% of variance in generic HRQoL (adjusted R^2=21%), $F_{9,164}$=6.09, P<.001. After adjusting for confounding factors, health literacy (b=.08, SE 0.02, 95% CI 0.04-0.12) was the only predictor variable positively associated with generic HRQoL (P<.001). Cohen f^2 for the effects of health literacy, eHealth literacy, and COPD knowledge on generic HRQOL was 0.12, which is a medium effect size.

Table 2. Hierarchical linear regression model examining effects of health literacy, electronic health literacy, and chronic obstructive pulmonary disease knowledge on generic health-related quality of life (centered at the mean).

Predictor	R^2 change	b	SE b	95% CI
Step 1	.16[a]			
Age[b]	—[c]	.01[d]	0.001	0.002 to 0.01
Gender (0=male, 1=female)	—	.02	0.03	−0.04 to 0.07
Race (0=nonwhite, 1=white)	—	−.01	0.02	−0.05 to 0.04
Marital status (0=unmarried, 1=married)	—	−.01	0.01	−0.03 to 0.01
Education (1=less than 8 years, 6=postgraduate)	—	.04[d]	0.01	0.02 to 0.07
Comorbidity (0=no, 1=yes)	—	−.06	0.03	−0.13 to 0.004
Step 2	.09[a]			
COPD[e] knowledge[a]	—	−.01	0.01	−0.03 to 0.002
Health literacy[a]	—	.08[a]	0.02	0.04 to 0.12
Electronic health literacy[a]	—	.02	0.02	−0.01 to 0.06
Total R^2	.25[a]	—	—	—
Adjusted R^2	.21[a]	—	—	—

[a]P<.001.

[b]Centered at the mean.

[c]Not applicable.

[d]P<.01.

[e]COPD: chronic obstructive pulmonary disease.

XSL•FO
RenderX

Table 3. Spearman rank correlations between EuroQoL-5D dimension scores and statistically significant independent variables emerging from hierarchical linear regression model predicting generic health-related quality of life.

EuroQol-5D dimensions	Age	Education level	Comorbidity (yes=1 or no=0)[a]	Health literacy	Electronic health literacy
(1) Mobility	.11 (.16)	.15 (.05)	−.78 (.43)	.20[b] (.01)	.15[c] (.049)
(2) Self-care	.12 (.13)	.18[c] (.02)	−1.66 (.10)	.42 (<.001)	.23[b] (.003)
(3) Usual activities	.10 (.20)	.22[b] (.004)	−.60 (.55)	.23[b] (.003)	.19[c] (.01)
(4) Pain/discomfort	.25[b] (.001)	.25[b] (.001)	−2.37[c] (.02)	.29 (<.001)	.05 (.49)
(5) Anxiety/depression	.33[d] (<.001)	.12 (.11)	−.46 (.64)	.22[b] (.004)	.11 (.16)

[a]Results of Mann-Whitney Wilcoxon test.

[b]P<.01.

[c]P<.05.

Table 3 lists correlations between EQ-5D dimension scores and statistically significant predictor variables emerging from the 2-stage hierarchical linear regression model predicting generic HRQoL. Age, education, and health literacy were all significantly (P<.05) associated with at least one EQ-5D dimension score. Older participants reported less difficulty managing pain/discomfort (ρ=.25, P=.01) and anxiety/depression (ρ=.33, P<.001). However, age was not significantly associated with mobility (ρ=.11, P=.16), self-care ability (ρ=.12, P=.13), or engagement in usual daily activities (ρ=.10, P=.20). Higher educational attainment was associated with better ability to wash or dress oneself (ρ=.18, P=.05), greater engagement in usual activities such as housework and family or leisure time (ρ=.22, P=.004), and less difficulty with pain and discomfort (ρ=.25, P=.001). Education level was not significantly associated with mobility or anxiety/depression, although the association between education and mobility approached statistical significance (ρ=.15, P=.05). Participants living without comorbidities reported significantly less pain and discomfort (U=−2.37, P=.02, η^2=0.03) but no significant associations with any other dimension of generic HRQoL. Higher health literacy scores were associated with significantly (P<.05) better scores on all 5 EQ-5D dimensions: mobility (ρ=.20, P=.01); self-care (ρ=.42, P<.001); usual activities (ρ=.23, P=.003); pain/discomfort (ρ=.29, P<.001); and anxiety/discomfort (ρ=.22, P=.004). eHealth literacy was positively associated with 3 generic HRQoL dimensions: mobility (ρ=.15, P=.049), self-care (ρ=.23, P=.003), and usual activities (ρ=.19, P=.01).

Test of Hypothesis 2: Health Literacy, Electronic Health Literacy, and Chronic Obstructive Pulmonary Disease Knowledge on Lung-Specific Health-Related Quality of Life

In the hierarchical regression model examining effects of health literacy, eHealth literacy, and COPD knowledge on lung-specific HRQoL, demographic characteristics and comorbidity status (step 1) accounted for 22.5% (adjusted R^2=19.7%) of the variance in lung-specific HRQoL, $F_{6,163}$=7.90, P<.001. Table 4 shows that older age (b=.21, SE 0.05, P<.001), higher education level (b=1.87, SE 0.47, P<.001), and living with no comorbidities (b=−2.85, SE 1.120, P<.05) were associated with better lung-specific HRQoL scores. Adding health literacy, eHealth literacy, and COPD knowledge as predictor variables (step 2) explained an additional 7.4% of variance in lung-specific HRQoL (adjusted R^2=26.4%), $F_{8,161}$=8.59, P<.001. After adjusting for potentially confounding factors, health literacy (b=2.63, SE 0.84, 95% CI 0.96-4.29, P<.001) and eHealth literacy (b=1.41, SE 0.67, 95% CI 0.09-2.73, P<.001) were positively associated with lung-specific HRQoL. Interestingly, after adjusting for confounding variables, COPD knowledge (b=−.56, SE 0.29, 95% CI −1.22 to −0.004, P<.05) was inversely associated with lung-specific HRQoL. In other words, higher COPD knowledge was related to lower lung-specific HRQoL among people living with COPD. Cohen f^2 for the effects of health literacy, eHealth literacy and COPD knowledge on lung specific HRQoL was 0.13, which is a medium effect size.

Table 5 describes correlations between CAT item scores and 6 statistically significant predictor variables emerging from the 2-stage hierarchical linear regression model. A total of 5 of the 6 predictor variables were significantly (P<.05) associated with at least three CAT item scores, except for COPD knowledge, which was not significantly associated with any CAT item. Increased age was associated with better symptom scores related to cough frequency (ρ=.23, P=.002), chest congestion (ρ=.21, P=.005), chest tightness (ρ=.30, P<.001), sleep quality (ρ=.32, P<.001), and energy level (ρ=.22, P=.003). Older participants also demonstrated greater confidence leaving home (ρ=.18, P=.015). Education level was positively associated (P<.05) with better scores on every CAT item, except for chest congestion (ρ=.10, P=.17). Participants without comorbidities reported less challenges with chest tightness (U=−2.46, P=.014, η^2=0.03), more confidence leaving home (U=−2.20, P=.03, η^2=0.03), and better energy levels (U=−2.90, P=.004, η^2=0.04). Health literacy was positively associated with scores on most lung-specific symptoms, including cough frequency (ρ=.18, P=.02), chest tightness (ρ=.25, P=.001), activity limitation at home (ρ=.24, P=.002), confidence leaving home (ρ=.29, P<.001), sleep quality (ρ=.25, P=.001), and energy level (ρ=.30, P<.001). eHealth literacy scores were also positively associated with over 60% (5/8) of the CAT items: chest congestion (ρ=.19, P=.013), chest tightness (ρ=.17, P=.02), activity limitations at home (ρ=.19, P=.01), sleep quality (ρ=.15, P=.04), and energy level (ρ=.18, P=.02).

XSL•FO
RenderX

Table 4. Hierarchical linear regression model examining effects of health literacy, electronic health literacy, and chronic obstructive pulmonary disease knowledge on lung-specific health-related quality of life (centered at the mean).

Predictor	R^2 change	*b*	SE *b*	95% CI
Step 1	.23[a]	—[b]	—	—
Age[c]	—	.21[a]	0.05	0.11 to 0.31
Gender (0=male, 1=female)	—	.72	0.98	−1.22 to 2.66
Race (0=nonwhite; 1=white)	—	−.64	0.82	−2.25 to 0.97
Marital status (0=unmarried, 1=married)	—	−.21	0.37	−0.95 to 0.53
Education (1=less than 8 years, 6=postgraduate)	—	1.87[a]	0.47	0.94 to 2.80
Comorbidity (0=no, 1=yes)	—	−2.85[e]	1.12	−5.21 to –0.49
Step 2	.09[a]	—	—	—
COPD[e] knowledge[c]	—	−.56[c]	0.28	−1.12 to −0.004
Health literacy[c]	—	2.63[f]	0.84	0.96 to 4.29
Electronic health literacy	—	1.41[f]	0.67	0.09 to 2.73
Total R^2	.32[a]	—	—	—
Adjusted R^2	.28[a]	—	—	—

[a]*P*<.001.

[b]Not applicable.

[c]Centered at the mean.

[d]*P*<.05.

[e]COPD: chronic obstructive pulmonary disease.

[f]*P*<.01.

Table 5. Spearman rank correlations between Chronic Obstructive Pulmonary Disease-Assessment Test Item scores and statistically significant predictor variables emerging from hierarchical linear regression models. Values in parentheses indicate *P* values.

CAT[a] items	Age	Education level	Comorbidity (yes=1/no=0)[b]	COPD[c] knowledge	Health literacy	Electronic health literacy
(1) Cough frequency	.23[d] (.002)	.20[d] (.009)	−1.18 (.24)	−.14 (.07)	.18[e] (.02)	.15 (.06)
(2) Chest congestion	.21[d] (.005)	.10 (.17)	−1.40 (.16)	−.05 (.51)	.15 (.06)	.19[d] (.013)
(3) Chest tightness	.30 (<.001)	.31 (<.001)	−2.46[e] (.014)	−.06 (.45)	.25[d] (.001)	.17[e] (.02)
(4) Breathlessness walking uphill	.10 (.19)	.23[d] (.002)	−.03 (.98)	.03 (.67)	.12 (.13)	.08 (.29)
(5) Activity limitations at home	.117 (.12)	.35 (<.001)	−1.83 (.07)	−.06 (.46)	.235[d] (.002)	.19[e] (.01)
(6) Confidence leaving home	.18[e] (.02)	.24[d] (.001)	−2.20[e](.03)	.04 (.62)	.29 (<.001)	.14 (.06)
(7) Sleep quality	.32 (<.001)	.24[d] (.001)	−1.71 (.08)	−.01 (.93)	.25[d] (.001)	.15[e] (.04)
(8) Energy level	.22[d] (.003)	.31 (<.001)	−2.90[d] (.004)	−.01 (.93)	.30 (<.001)	.18[e] (.02)

[a]CAT: COPD Assessment Test.

[b]Results of Mann-Whitney Wilcoxon test.

[c]COPD: chronic obstructive pulmonary disease.

[d]*P*<.01.

[e]*P*<.05.

XSL•FO
RenderX

Table 6. Moderating effect of chronic obstructive pulmonary disease knowledge on the relationship between health literacy and lung-specific health-related quality of life (centered at the mean).

Predictor	R^2 change	*b*	SE *b*	95% CI
Step 1	.23[a]	—[b]	—	—
Age[c]	—	0.21[a]	0.05	0.11 to 0.31
Gender (0=male, 1=female)	—	0.72	0.98	−1.22 to 2.66
Race (0=nonwhite, 1=white)	—	−0.64	0.82	−2.25 to 0.97
Marital status (0=unmarried, 1=married)	—	−0.21	0.37	−0.95 to 0.053
Education (1=less than 8 years, 6=postgraduate)	—	1.87[a]	0.47	0.94 to 2.80
Comorbidity (0=no, 1=yes)	—	–2.85[e]	1.20	–5.21 to –0.49
Step 2	.07[a]	—	—	—
COPD[e] knowledge[c]	—	−0.45	0.28	−1.00 to 0.11
Health literacy[c]	—	2.63[f]	0.84	0.96 to 4.29
Step 3	.02 (*P*=.055)	—	—	—
COPD knowledge health literacy	—	0.89	0.46	–0.02 to 1.80
Total R^2	.31[a]	—	—	—
Adjusted R^2	.27[a]	—	—	—

[a]*P*<.001.

[b]Not applicable.

[c]Centered at the mean.

[d]*P*<.05.

[e]COPD: chronic obstructive pulmonary disease.

[f]*P*<.01.

Test of Hypothesis 3: Moderating Effect of Chronic Obstructive Pulmonary Disease Knowledge on the Relationship Between Health Literacy and Lung-Specific Health-Related Quality of Life

The interaction between COPD knowledge and health literacy (step 3) did not account for significantly more variance in lung-specific HRQoL than COPD knowledge and health literacy alone, R^2 change=.01, *P*=.10 (Table 6). Moreover, the interaction effect of health literacy and lung-specific HRQoL (*b*=.890, SE 0.461) only approached statistical significance (*P*=.055). Greater knowledge of COPD did not significantly increase (or enhance) the effect of health literacy on lung-specific HRQoL in this sample of patients living with COPD. Cohen f^2 for the moderating effects of COPD knowledge on health literacy and lung-specific HRQoL was 0.12, which is a medium effect size.

Test of Hypothesis 4: Moderating Effect of Chronic Obstructive Pulmonary Disease Knowledge on the Relationship Between Electronic Health Literacy and Lung-Specific Health-Related Quality of Life

The interaction between COPD knowledge and eHealth literacy (step 3) also did not account for significantly more variance in lung-specific HRQoL than COPD knowledge and health literacy did alone, R^2 change=.01, *P*=.28 (Table 7). Greater knowledge of COPD did not significantly increase (or enhance) the effect of eHealth literacy on lung-specific HRQoL in the sample of patients living with COPD. Cohen f^2 for the moderating effects of COPD knowledge on eHealth literacy and lung-specific HRQoL was 0.07, which is a small effect size.

XSL•FO
RenderX

Table 7. Moderating effect of chronic obstructive pulmonary disease knowledge on the relationship between electronic health literacy and lung-specific health-related quality of life.

Predictor	R^2 change	*b*	SE *b*	95% CI
Step 1	.23[a]	—[b]	—	—
Age[c]	—	0.21[a]	0.05	0.11 to 0.31
Gender (0=male, 1=female).	—	0.72	0.98	−1.22 to 2.66
Race (0=nonwhite; 1=white)	—	−0.64	0.82	−2.25 to 0.97
Marital status (0=unmarried, 1=married)	—	−0.21	0.37	−0.95 to 0.53
Education (1=less than 8 years, 6=postgraduate)	—	1.87[a]	0.47	0.94 to 2.80
Comorbidity (0=no, 1=yes)	—	−2.85[d]	1.20	−5.21 to –0.49
Step 2	.05[e]	—	—	—
COPD[f] knowledge[c]	—	−0.56	0.29	−1.13 to 0.02
Electronic health literacy[c]	—	1.99[e]	0.66	0.69 to 3.29
Step 3	.002	—	—	—
COPD knowledge electronic health literacy	—	−.025	0.40	−1.04 to 0.54
Total R^2	.28[a]	—	—	—
Adjusted R^2	.24[a]	—	—	—

[a]*P*<.001.

[b]Not applicable.

[c]Centered at the mean.

[d]*P*<.05.

[e]*P*<.01.

[f]COPD: chronic obstructive pulmonary disease.

Discussion

Principal Findings

This study aimed to assess how health literacy, COPD-related knowledge, and eHealth literacy are associated with both generic HRQoL and lung-specific HRQoL for people diagnosed with COPD while controlling for the effect of comorbid conditions, socioeconomic status, education level, and other factors known to be associated with HRQoL. Overall, results provide support for the importance of considering both health and eHealth literacy levels when developing patient education and self-management support programs for people living with COPD. More specifically, results revealed that health literacy was significantly associated with all 5 EQ-5D dimensions measuring generic HRQoL, whereas eHealth literacy was associated with significantly better scores on 3 generic HRQoL dimensions (mobility, self-care, and usual activities). Both health and eHealth literacies were positively associated with lung-specific HRQoL; however, higher levels of COPD knowledge did not significantly enhance the impact of health literacy or eHealth literacy on lung-specific HRQoL. These results also signified approximately moderate ($f^2 \geq 0.15$) effect sizes according to Cohen's guidelines [49], specifically with the effects of health literacy and eHealth literacy on both generic and lung-specific HRQoL yielding effect size coefficients in the moderate/medium range. Results indicate the need to consider both eHealth literacy and health literacy levels of patients when developing patient education and support programs, particularly when the intended outcome is to enhance lung-specific quality of life. In addition, given that COPD knowledge did not demonstrate an impact on health literacy and eHealth literacy, future research should explore addressing these 2 constructs in the context of patient self-management programs rather than traditional patient education programs that solely focus on knowledge dissemination alone.

Influence of Health Literacy, Electronic Health Literacy, and Chronic Obstructive Pulmonary Disease Knowledge on Generic Health-Related Quality of Life

Overall, results showed similar HRQoL mean scores and correlates (ie, age, education level) to previous studies. HRQoL is generally lower for individuals living with COPD compared with the general population [50]. This study found that the mean EQ-5D score for measuring generic HRQoL among participants with COPD was .723 (SD 0.18) on a 0 to 1 scale, with the higher score meaning better health. In a report of nationally representative values for the US adult population, the average EQ-5D score for both males and females is reported to be above 0.80 for most age brackets, except for adults 80 years and older, which is slightly lower, but above the mean EQ-5D score for participants in this study [51]. The significant associations with age and education level to at least one dimension as assessed here by the EQ-5D are consistent with previous studies that

XSL•FO
RenderX

have demonstrated that age and education influence HRQoL [52-54]. According to van Manen and colleagues [54], HRQoL is a multidimensional construct that encompasses both a physical and mental component. Age negatively affects general health and physical function, whereas it positively influences mental health; and in general, those with higher, formal education levels report better mental health [54]. In addition, education level has been found to be a significant predictor of poor health outcomes with COPD [55] while directly affecting health status through associations with patient ability to navigate a complex health care system [56,57].

However, as this study showed, education and age were not significantly associated with all dimensions of HRQoL, as measured by the EQ-5D; therefore, more research should be conducted to better explain which dimensions are influenced by these variables and in what ways. Furthermore, after adjusting for confounding variables, including age and education level, only health literacy was found to be significantly associated with generic HRQoL. These findings are consistent with 2 previous research studies [22,24] that indicate limited health literacy influences general HRQoL of people living with COPD. Specifically, these studies indicate that lower health literacy levels are associated with worse COPD severity and greater levels of helplessness reported among patients living with COPD [22,24]. In addition, these research results suggest that the role of health literacy in COPD may be more important than education when it comes to health outcomes [22].

Although the current investigation demonstrates a strong association between health literacy and general HRQoL, even when controlling for sociodemographic variables such as education and age, eHealth literacy and COPD knowledge were not statistically associated with all 5 EQ-5D dimensions. However, eHealth literacy was significantly associated with scores on 3 generic HRQoL dimensions, including mobility, usual activities, and self-care. Collectively, these dimensions reflect the ability of a patient to manage their illness, and prior research supports that higher levels of eHealth literacy can improve patient self-management of COPD [26]. Nevertheless, there are few existing studies examining associations between eHealth literacy with self-reported health status, which limits what inferences can be made. More longitudinal studies are needed to examine whether better self-management and HRQoL outcomes are experienced by patients with COPD who are more eHealth literate [58].

The results regarding COPD knowledge support previous research findings that the transfer of disease-specific knowledge does not necessarily have a long-term impact on behavior change that could lead to improved quality of life for patients living with COPD [59]. In fact, these findings imply that COPD knowledge alone is not enough to improve HRQoL; however, the potential for using COPD knowledge in conjunction with skill acquisition approaches and self-efficacy–building strategies, with a goal of reaching behavior change, could prove to be collective core elements that are helpful in achieving long-term impact on general HRQoL of patients with COPD [59]. More research is warranted to assess the specific role COPD knowledge plays in factors impacting behavior changes that are associated with general HRQoL.

Influence of Health Literacy, Electronic Health Literacy, and Chronic Obstructive Pulmonary Disease Knowledge on Lung-Specific Health-Related Quality of Life

The results, overall, showed that after adjusting for potentially confounding variables, health literacy and eHealth literacy levels were significantly associated with lung-specific HRQoL. The impact of distressing COPD symptoms, such as dyspnea, fatigue, coughing, anxiety, wheezing, pain interference, and breathlessness, can help explain much of the variance in lung-specific HRQoL among patients living with the disease [60-62]. This study sought to evaluate the impact of health literacy, eHealth literacy, and COPD knowledge on lung-specific HRQoL. These findings support previous research on the associations between poor health literacy and COPD-related outcomes, revealing lower levels of health literacy relate to worse COPD severity, higher levels of COPD helplessness, and poorer levels of lung-specific HRQoL [22]. As such, this study demonstrates similar results to that of past research indicating poor health literacy is a risk factor associated with poor health outcomes for patients living with COPD, including lung-specific HRQoL, and these patients could be more at-risk for COPD-related emergency health care utilization and overall economic cost burdens associated with the disease. Certainly, these results, as they stand, do not provide reason to support causal relationships between health literacy and lower health outcomes, including lung-specific HRQoL; however, Omachi and colleagues [22] suggest that such relationships could exist, and more research on causality is needed. In addition, the results support the need to consider addressing both health literacy and eHealth literacy in COPD patient education programming.

eHealth literacy is also positively associated with patients' lung-specific HRQoL in this sample. The eHealth literacy scores are comparable with those in other studies assessing eHealth literacy in the general older adult population [63], as a majority of adults self-reported moderate to above average levels of eHealth literacy. Patients living with COPD who can evaluate the quality of Web-based health information for use in their own care are more amenable to engaging in formal disease self-management programs and more skilled in making decisions related to their specific diagnosis [64]. Therefore, individuals with COPD who have these skills to delineate the quality of resources found on the internet for clinical decision making to manage the distressing symptoms of COPD have enhanced lung-specific HRQoL. Future patient-centered research should emphasize the importance of patients having better eHealth literacy skills to not only enhance lung-specific HRQoL but also potentially better health outcomes among patients living with COPD.

Although eHealth literacy was positively associated with lung-specific HRQoL, it was not associated with generic HRQoL. Patients with COPD who use Web-based media for health-related purposes readily disclose unmet needs related to respiratory symptoms, including cough, mucus production, and dyspnea [65]. Rather than accessing generic health information, patients who go on the internet may be more inclined to access health information that is relevant to COPD and its

XSL•FO
RenderX

symptomology [29]. This includes self-management topics focused more on lung-specific outcomes (eg, shortness of breath and chest congestion) rather than generic HRQoL outcomes (eg, mobility, pain, or discomfort). Therefore, high eHealth literate patients with COPD may be accessing self-management information about behavior changes that have the potential to improve lung functioning yet not directly lead to enhanced general HRQoL. More research is needed to understand eHealth approaches to self-management education for individuals with COPD, and how eHealth skill-building strategies impact overall, functional health outcomes for patients.

Effect of Chronic Obstructive Pulmonary Disease Knowledge on the Relationship Between Health Literacy, Electronic Health Literacy, and Generic/Lung-Specific Health-Related Quality of Life

Greater COPD knowledge did not enhance the effect of health literacy or eHealth literacy on general or lung-specific HRQoL. COPD knowledge, however, was inversely related to lung-specific HRQoL; therefore, the higher the knowledge level, the lower was the self-reported lung-specific HRQoL of study participants. This finding was not anticipated, but one explanation for this result is that even with increased knowledge of the disease, patients might lack the skills or self-efficacy to change their behaviors that are linked to one's disease state, impacting their lung-specific HRQoL. In addition, it is possible that more knowledge related to the serious diagnosis may result in higher levels of anxiety, fear, and even depression, which could, in turn, negatively affect their general and/or lung-specific HRQoL [66].

Limitations

This study has several limitations related to recruitment techniques employed, a low sample size, and lack of a representative sample. Despite collaborating with the COPD Foundation's National Research Registry, this Web-based survey had a low response rate of patients with COPD (13.7%) as compared with other internet surveys conducted with the general population [67]. Our recruitment approach was limited to a single email. Given that recruitment approaches are rarely effective at reaching target enrollment with a one-time invitation, our 13.7% response rate is quite commendable for this patient population. Patients with COPD are generally older and experience frequent hospitalizations that prevent them from engaging in their everyday routines, which may include email checking. This study recruited from a research registry, which requires patients to register for program, and may not be representative of all patients with COPD and may have contributed to the low response rate. This selection bias is essentially unavoidable when using research registries but should be considered when interpreting results. In addition, the magnitude of this bias toward the main results is unknown, as data on unenrolled patients from the underlying population base are unavailable for further analyses. It would only help internal validity to have data on patients not enrolled in the national registry on the exact factors being assessed in this study. Therefore, generalizability of results may be affected and limited to patients in similar registries. Future studies should explore if results from this study are similar with recruitment that includes both Web-based and non-Web-based techniques.

The statistically significant associations found in this cross-sectional study cannot be used to establish cause-and-effect relationships; however, use of hierarchical predictor variable entry provided some advantages when attempting to identify relationships between health/eHealth literacy, COPD knowledge, and relevant health-related outcomes. The construct of eHealth literacy was measured with eHEALS, a widely used scale that has been questioned in the age of new media where communication technologies such as social media are pervasive [68,69]. Studies assessing population eHealth literacy may consider using more updated measures now available to researchers [68,70].

In addition, low socioeconomic status minorities living with COPD were underrepresented. Most participants in this study identified as being white and highly educated, which is not reflective of the sociodemographics of COPD [71]. As mentioned elsewhere [26], future studies assessing health information technology use among patients with COPD should seek to recruit larger numbers of participants with more diverse racial or ethnic, geographic, and economic characteristics. Recruitment should also entail more inclusive strategies than email only, which will likely encourage participation among patients being treated in community-based and clinical settings.

Conclusions

Although previous research has shown a relatively high prevalence of low health literacy among individuals living with COPD, little empirical attention has been directed at exploring the cognitive and health literacy–related skills that can influence patients' HRQoL. Findings from this study indicated that health literacy, but not eHealth literacy, was positively associated with generic HRQoL. In addition, health literacy and eHealth literacy were found to be positively associated with lung-specific HRQoL, with higher COPD knowledge associated with lower lung-specific HRQoL. These results confirm the importance of considering patient health and eHealth literacy levels when designing self-management support programs for people living with COPD. Future research should explore the impact of delivering COPD interventions aimed at improving eHealth and health literacy, particularly when self-management program goals include improving HRQoL outcomes.

Acknowledgments

The authors would like to thank the representatives from the COPD Foundation's COPD Research Registry for their assistance with study recruitment.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Martinez CH, Mannino DM, Jaimes FA, Curtis JL, Han MK, Hansel NN, et al. ndiagnosed obstructive lung disease in the United States. Associated factors and long-term mortality. Ann Am Thorac Soc 2015 Dec;12(12):1788-1795 [FREE Full text] [doi: 10.1513/AnnalsATS.201506-388OC] [Medline: 26524488]
2. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 2013 Feb 15;187(4):347-365. [doi: 10.1164/rccm.201204-0596PP] [Medline: 22878278]
3. World Health Organization. Burden of COPD URL: https://www.who.int/respiratory/copd/burden/en/ [accessed 2019-05-16] [WebCite Cache ID 78PqEgePj]
4. Obi J, Mehari A, Gillum R. Mortality related to chronic obstructive pulmonary disease and co-morbidities in the United States, a multiple causes of death analysis. COPD 2018 Apr;15(2):200-205. [doi: 10.1080/15412555.2018.1454897] [Medline: 29697272]
5. Gardiner C, Gott M, Payne S, Small N, Barnes S, Halpin D, et al. Exploring the care needs of patients with advanced COPD: an overview of the literature. Respir Med 2010 Feb;104(2):159-165 [FREE Full text] [doi: 10.1016/j.rmed.2009.09.007] [Medline: 19818590]
6. Shah T, Churpek MM, Coca PM, Konetzka RT. Understanding why patients with COPD get readmitted: a large national study to delineate the Medicare population for the readmissions penalty expansion. Chest 2015 May;147(5):1219-1226 [FREE Full text] [doi: 10.1378/chest.14-2181] [Medline: 25539483]
7. Sharif R, Parekh TM, Pierson KS, Kuo Y, Sharma G. Predictors of early readmission among patients 40 to 64 years of age hospitalized for chronic obstructive pulmonary disease. Ann Am Thorac Soc 2014 Jun;11(5):685-694 [FREE Full text] [doi: 10.1513/AnnalsATS.201310-358OC] [Medline: 24784958]
8. Sharma G, Kuo Y, Freeman JL, Zhang DD, Goodwin JS. Outpatient follow-up visit and 30-day emergency department visit and readmission in patients hospitalized for chronic obstructive pulmonary disease. Arch Intern Med 2010 Oct 11;170(18):1664-1670 [FREE Full text] [doi: 10.1001/archinternmed.2010.345] [Medline: 20937926]
9. Makaryus AN, Friedman EA. Patients' understanding of their treatment plans and diagnosis at discharge. Mayo Clin Proc 2005 Aug;80(8):991-994. [doi: 10.4065/80.8.991] [Medline: 16092576]
10. Rigual J, Centanaro G, Pelaez F, Hermosa J, Forcén E, Álvarez-Sala J, et al. Factors associated with therapeutic compliance in COPD. Eur Respir J 2017;50(suppl 61):3881. [doi: 10.1183/1393003.congress-2017.PA3881]
11. Claxton AJ, Cramer J, Pierce C. A systematic review of the associations between dose regimens and medication compliance. Clin Ther 2001 Aug;23(8):1296-1310. [Medline: 11558866]
12. Hernandez P, Balter M, Bourbeau J, Hodder R. Living with chronic obstructive pulmonary disease: a survey of patients' knowledge and attitudes. Respir Med 2009 Jul;103(7):1004-1012 [FREE Full text] [doi: 10.1016/j.rmed.2009.01.018] [Medline: 19269150]
13. Martinez CH, Raparla S, Plauschinat CA, Giardino ND, Rogers B, Beresford J, et al. Gender differences in symptoms and care delivery for chronic obstructive pulmonary disease. J Womens Health (Larchmt) 2012 Dec;21(12):1267-1274 [FREE Full text] [doi: 10.1089/jwh.2012.3650] [Medline: 23210491]
14. Jones I, Kirby A, Ormiston P, Loomba Y, Chan K, Rout J, et al. The needs of patients dying of chronic obstructive pulmonary disease in the community. Fam Pract 2004 Jun;21(3):310-313. [Medline: 15128695]
15. Martinez CH, Diaz AA, Parulekar AD, Rennard SI, Kanner RE, Hansel NN, COPDGeneSPIROMICS Investigators. Age-related differences in health-related quality of life in COPD: an analysis of the COPDGene and SPIROMICS cohorts. Chest 2016 Apr;149(4):927-935 [FREE Full text] [doi: 10.1016/j.chest.2015.11.025] [Medline: 26836895]
16. Burgel P, Escamilla R, Perez T, Carré P, Caillaud D, Chanez P, INITIATIVES BPCO Scientific Committee. Impact of comorbidities on COPD-specific health-related quality of life. Respir Med 2013 Feb;107(2):233-241 [FREE Full text] [doi: 10.1016/j.rmed.2012.10.002] [Medline: 23098687]
17. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2095-2128. [doi: 10.1016/S0140-6736(12)61728-0] [Medline: 23245604]
18. Putcha N, Puhan MA, Drummond MB, Han MK, Regan EA, Hanania NA, et al. A simplified score to quantify comorbidity in COPD. PLoS One 2014 Dec;9(12):e114438 [FREE Full text] [doi: 10.1371/journal.pone.0114438] [Medline: 25514500]
19. Zhang Q, Liao J, Liao X, Wu X, Wan M, Wang C, et al. Disease knowledge level is a noteworthy risk factor of anxiety and depression in patients with chronic obstructive pulmonary disease: a cross-sectional study. BMC Pulm Med 2014 May 28;14:92 [FREE Full text] [doi: 10.1186/1471-2466-14-92] [Medline: 24884452]
20. COPD Foundation. COPD Foundation releases groundbreaking COPE survey results: Low patient awareness about COPD exacerbations poses URL: https://www.copdfoundation.org/About-Us/Press-Room/Press-Releases/Article/883/

XSL•FO
RenderX

COPD-Foundation-Releases-Groundbreaking-COPE-Survey-Results-Low-Patient-Awarene.aspx [accessed 2018-07-11] [WebCite Cache ID 70qIQr4Gd]
21. Nakken N, Janssen DJ, van den Bogaart EH, Muris JW, Vercoulen JH, Custers FL, et al. Knowledge gaps in patients with COPD and their proxies. BMC Pulm Med 2017 Oct 30;17(1):136 [FREE Full text] [doi: 10.1186/s12890-017-0481-8] [Medline: 29084521]
22. Omachi TA, Sarkar U, Yelin EH, Blanc PD, Katz PP. Lower health literacy is associated with poorer health status and outcomes in chronic obstructive pulmonary disease. J Gen Intern Med 2013 Jan;28(1):74-81 [FREE Full text] [doi: 10.1007/s11606-012-2177-3] [Medline: 22890622]
23. Sadeghi S, Brooks D, Stagg-Peterson S, Goldstein R. Growing awareness of the importance of health literacy in individuals with COPD. COPD 2013 Feb;10(1):72-78. [doi: 10.3109/15412555.2012.727919] [Medline: 23413895]
24. Puente-Maestu L, Calle M, Rodríguez-Hermosa JL, Campuzano A, de Miguel Díez J, Álvarez-Sala JL, et al. Health literacy and health outcomes in chronic obstructive pulmonary disease. Respir Med 2016 Dec;115:78-82 [FREE Full text] [doi: 10.1016/j.rmed.2016.04.016] [Medline: 27215508]
25. Kale MS, Federman AD, Krauskopf K, Wolf M, O'Conor R, Martynenko M, et al. The association of health literacy with illness and medication beliefs among patients with chronic obstructive pulmonary disease. PLoS One 2015 Apr;10(4):e0123937 [FREE Full text] [doi: 10.1371/journal.pone.0123937] [Medline: 25915420]
26. Stellefson ML, Shuster JJ, Chaney BH, Paige SR, Alber JM, Chaney JD, et al. Web-based Health Information Seeking and eHealth Literacy among Patients Living with Chronic Obstructive Pulmonary Disease (COPD). Health Commun 2018 Dec 5;33(12):1410-1424. [doi: 10.1080/10410236.2017.1353868] [Medline: 28872905]
27. Roberts NJ, Ghiassi R, Partridge MR. Health literacy in COPD. Int J Chron Obstruct Pulmon Dis 2008;3(4):499-507 [FREE Full text] [Medline: 19281068]
28. Slatore CG, Cecere LM, Reinke LF, Ganzini L, Udris EM, Moss BR, et al. Patient-clinician communication: associations with important health outcomes among veterans with COPD. Chest 2010 Sep;138(3):628-634 [FREE Full text] [doi: 10.1378/chest.09-2328] [Medline: 20299633]
29. Martinez CH, St JBL, Plauschinat CA, Rogers B, Beresford J, Martinez FJ, et al. Internet access and use by COPD patients in the National Emphysema/COPD Association Survey. BMC Pulm Med 2014;14:66 [FREE Full text] [doi: 10.1186/1471-2466-14-66] [Medline: 24755090]
30. Borycki E. M-health: can chronic obstructive pulmonary disease patients use mobile phones and associated software to self-manage their disease? Stud Health Technol Inform 2012;172:79-84. [Medline: 22910504]
31. Hopkinson NS. Social media as a source of information for patients with chronic obstructive pulmonary disease. Chron Respir Dis 2014 May;11(2):59-60. [doi: 10.1177/1479972314528959] [Medline: 24728655]
32. Stellefson M, Chaney B, Ochipa K, Chaney D, Haider Z, Hanik B, et al. YouTube as a source of chronic obstructive pulmonary disease patient education: a social media content analysis. Chron Respir Dis 2014 May;11(2):61-71 [FREE Full text] [doi: 10.1177/1479972314525058] [Medline: 24659212]
33. National Heart, Lung, and Blood Institute. COPD: The more you know, the better for you and your loved ones URL: https://www.nhlbi.nih.gov/files/docs/public/lung/copd-atrisk.pdf [WebCite Cache ID 75oqjjonu]
34. Maples P, Franks A, Ray S, Stevens AB, Wallace LS. Development and validation of a low-literacy chronic obstructive pulmonary disease knowledge questionnaire (COPD-Q). Patient Educ Couns 2010 Oct;81(1):19-22. [doi: 10.1016/j.pec.2009.11.020] [Medline: 20044232]
35. Ray SM, Helmer RS, Stevens AB, Franks AS, Wallace LS. Clinical utility of the chronic obstructive pulmonary disease knowledge questionnaire. Fam Med 2013 Mar;45(3):197-200 [FREE Full text] [Medline: 23463434]
36. Puente-Maestu L, Chancafe-Morgan J, Calle M, Rodríguez-Hermosa JL, Malo DM, Ortega-González A, en nombre del grupo GemEPOC. Validation of the Spanish Version of the COPD-Q Questionnaire on COPD Knowledge. Arch Bronconeumol 2016 Jan;52(1):12-16 [FREE Full text] [doi: 10.1016/j.arbres.2015.03.013] [Medline: 26026685]
37. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
38. Haun JN, Valerio MA, McCormack LA, Sørensen K, Paasche-Orlow MK. Health literacy measurement: an inventory and descriptive summary of 51 instruments. J Health Commun 2014;19 Suppl 2:302-333. [doi: 10.1080/10810730.2014.936571] [Medline: 25315600]
39. Wallace LS, Rogers ES, Roskos SE, Holiday DB, Weiss BD. Brief report: screening items to identify patients with limited health literacy skills. J Gen Intern Med 2006 Aug;21(8):874-877. [doi: 10.1111/j.1525-1497.2006.00532.x] [Medline: 16881950]
40. Chew LD, Griffin JM, Partin MR, Noorbaloochi S, Grill JP, Snyder A, et al. Validation of screening questions for limited health literacy in a large VA outpatient population. J Gen Intern Med 2008 May;23(5):561-566 [FREE Full text] [doi: 10.1007/s11606-008-0520-5] [Medline: 18335281]
41. 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]

XSL•FO
RenderX

42. Collins SA, Currie LM, Bakken S, Vawdrey DK, Stone PW. Health literacy screening instruments for eHealth applications: a systematic review. J Biomed Inform 2012 Jun;45(3):598-607 [FREE Full text] [doi: 10.1016/j.jbi.2012.04.001] [Medline: 22521719]
43. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
44. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care 2005 Mar;43(3):203-220. [doi: 10.1097/00005650-200503000-00003] [Medline: 15725977]
45. Jones PW, Harding G, Berry P, Wiklund I, Chen W, Kline LN. Development and first validation of the COPD Assessment Test. Eur Respir J 2009 Sep;34(3):648-654 [FREE Full text] [doi: 10.1183/09031936.00102509] [Medline: 19720809]
46. Dodd JW, Hogg L, Nolan J, Jefford H, Grant A, Lord VM, et al. The COPD assessment test (CAT): response to pulmonary rehabilitation. A multicentre, prospective study. Thorax 2011 May;66(5):425-429. [doi: 10.1136/thx.2010.156372] [Medline: 21398686]
47. Cohen J, Cohen P, West S, Aiken L. Applied Multiple Regression/Correlation Analysis for Behavioral Sciences. Third Edition. New York, NY: Routledge; 2003.
48. Selya AS, Rose J, Dierker L, Hedeker D, Mermelstein R. A practical guide to calculating Cohen's f(2), a measure of local effect size, from PROC MIXED. Front Psychol 2012;3:111 [FREE Full text] [doi: 10.3389/fpsyg.2012.00111] [Medline: 22529829]
49. Cohen J. Statistical Power Analysis For The Behavioral Sciences (Second Edition). Hillsdale, NJ: Routledge; 1988.
50. Brandl M, Böhmer MM, Brandstetter S, Finger T, Fischer W, Pfeifer M, et al. Factors associated with generic health-related quality of life (HRQOL) in patients with chronic obstructive pulmonary disease (COPD): a cross-sectional study. J Thorac Dis 2018 Feb;10(2):766-775 [FREE Full text] [doi: 10.21037/jtd.2018.01.122] [Medline: 29607147]
51. Hanmer J, Lawrence WF, Anderson JP, Kaplan RM, Fryback DG. Report of nationally representative values for the noninstitutionalized US adult population for 7 health-related quality-of-life scores. Med Decis Making 2006;26(4):391-400. [doi: 10.1177/0272989X06290497] [Medline: 16855127]
52. Ferrer M, Alonso J, Morera J, Marrades RM, Khalaf A, Aguar MC, et al. Chronic obstructive pulmonary disease stage and health-related quality of life. The Quality of Life of Chronic Obstructive Pulmonary Disease Study Group. Ann Intern Med 1997 Dec 15;127(12):1072-1079. [Medline: 9412309]
53. Carrasco GP, de Miguel Díez J, Rejas GJ, Centeno AM, Gobartt VE, Gil DM, et al. Negative impact of chronic obstructive pulmonary disease on the health-related quality of life of patients. Results of the EPIDEPOC study. Health Qual Life Outcomes 2006 May 23;4:31 [FREE Full text] [doi: 10.1186/1477-7525-4-31] [Medline: 16719899]
54. van Manen MJG, Bindels PJE, Dekker FW, Bottema BJ, van der Zee JS, Ijzermans CJ, et al. The influence of COPD on health-related quality of life independent of the influence of comorbidity. J Clin Epidemiol 2003 Dec;56(12):1177-1184. [Medline: 14680668]
55. Eisner MD, Blanc PD, Omachi TA, Yelin EH, Sidney S, Katz PP, et al. Socioeconomic status, race and COPD health outcomes. J Epidemiol Community Health 2011 Jan;65(1):26-34 [FREE Full text] [doi: 10.1136/jech.2009.089722] [Medline: 19854747]
56. Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. Arch Intern Med 2007 Jul 23;167(14):1503-1509. [doi: 10.1001/archinte.167.14.1503] [Medline: 17646604]
57. Blane D. Commentary: explanations of the difference in mortality risk between different educational groups. Int J Epidemiol 2003 Jun;32(3):355-356. [doi: 10.1093/ije/dyg078] [Medline: 12777419]
58. Neter E, Brainin E. Association between health literacy, eHealth literacy, and health outcomes among patients with long-term conditions: a systematic review. Eur Psychol 2019 Jan;24(1):68-81. [doi: 10.1027/1016-9040/a000350]
59. Monninkhof E, van der Valk P, van der Palen J, van Herwaarden C, Partridge MR, Zielhuis G. Self-management education for patients with chronic obstructive pulmonary disease: a systematic review. Thorax 2003 May;58(5):394-398 [FREE Full text] [Medline: 12728158]
60. Kapella MC, Larson JL, Patel MK, Covey MK, Berry JK. Subjective fatigue, influencing variables, and consequences in chronic obstructive pulmonary disease. Nurs Res 2006;55(1):10-17. [Medline: 16439924]
61. Borge CR, Wahl AK, Moum T. Association of breathlessness with multiple symptoms in chronic obstructive pulmonary disease. J Adv Nurs 2010 Dec;66(12):2688-2700. [doi: 10.1111/j.1365-2648.2010.05447.x] [Medline: 20825511]
62. Borge CR, Wahl AK, Moum T. Pain and quality of life with chronic obstructive pulmonary disease. Heart Lung 2011;40(3):e90-101. [doi: 10.1016/j.hrtlng.2010.10.009] [Medline: 21444112]
63. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
64. Clarke MA, Moore JL, Steege LM, Koopman RJ, Belden JL, Canfield SM, et al. Health information needs, sources, and barriers of primary care patients to achieve patient-centered care: a literature review. Health Informatics J 2015 Sep 15:992-1016. [doi: 10.1177/1460458215602939] [Medline: 26377952]

XSL•FO
RenderX

65. Cook NS, Kostikas K, Gruenberger J, Shah B, Pathak P, Kaur VP, et al. Patients' perspectives on COPD: findings from a social media listening study. ERJ Open Res 2019 Feb;5(1):00128-02018. [doi: 10.1183/23120541.00128-2018] [Medline: 30775374]
66. Pumar MI, Gray CR, Walsh JR, Yang IA, Rolls TA, Ward DL. Anxiety and depression-Important psychological comorbidities of COPD. J Thorac Dis 2014 Nov;6(11):1615-1631 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2014.09.28] [Medline: 25478202]
67. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? Assess Eval Higher Edu 2008 Jun;33(3):301-314. [doi: 10.1080/02602930701293231]
68. Paige SR, Miller MD, Krieger JL, Stellefson M, Cheong J. Electronic health literacy across the lifespan: measurement invariance study. J Med Internet Res 2018 Jul 9;20(7):e10434 [FREE Full text] [doi: 10.2196/10434] [Medline: 29986848]
69. van der Vaart R, van Deursen AJ, Drossaert CH, Taal E, van Dijk JA, van de Laar MA. Does the eHealth Literacy Scale (eHEALS) measure what it intends to measure? Validation of a Dutch version of the eHEALS in two adult populations. J Med Internet Res 2011;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
70. Stellefson M, Paige SR, Tennant B, Alber JM, Chaney BH, Chaney D, et al. Reliability and validity of the telephone-based eHealth literacy scale among older adults: cross-sectional survey. J Med Internet Res 2017 Oct 26;19(10):e362 [FREE Full text] [doi: 10.2196/jmir.8481] [Medline: 29074471]
71. Gershon AS, Dolmage TE, Stephenson A, Jackson B. Chronic obstructive pulmonary disease and socioeconomic status: a systematic review. COPD 2012 Jun;9(3):216-226. [doi: 10.3109/15412555.2011.648030] [Medline: 22497534]

Abbreviations

CAT: COPD Assessment Test
COPD: chronic obstructive pulmonary disease
COPD-Q: Chronic Obstructive Pulmonary Disease Knowledge Questionnaire
eHealth: electronic health
eHEALS: eHealth Literacy Scale
EQ-5D: EuroQol-5D
HLSQ: Health Literacy Screening Questionnaire
HRQoL: health-related quality of life
IRB: institutional review board
VIF: variance inflation factor

Edited by G Eysenbach; submitted 11.09.18; peer-reviewed by S Mitsutake, WC Hsu; comments to author 07.01.19; revised version received 01.02.19; accepted 12.04.19; published 06.06.19.

Please cite as:
Stellefson M, Paige SR, Alber JM, Chaney BH, Chaney D, Apperson A, Mohan A
Association Between Health Literacy, Electronic Health Literacy, Disease-Specific Knowledge, and Health-Related Quality of Life Among Adults With Chronic Obstructive Pulmonary Disease: Cross-Sectional Study
J Med Internet Res 2019;21(6):e12165
URL: https://www.jmir.org/2019/6/e12165/
doi:10.2196/12165
PMID:31172962

©Michael Stellefson, Samantha R Paige, Julia M Alber, Beth H Chaney, Don Chaney, Avery Apperson, Arjun Mohan. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Visibility of Community Nursing Within an Administrative Health Classification System: Evaluation of Content Coverage

Lorraine J Block[1], RN, MSN; Leanne M Currie[1], RN, PhD; Nicholas R Hardiker[2], RN, PhD, FACMI; Gillian Strudwick[3], RN, PhD

[1]School of Nursing, University of British Columbia, Vancouver, BC, Canada

[2]School of Human and Health Sciences, University of Huddersfield, Huddersfield, United Kingdom

[3]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
Lorraine J Block, RN, MSN
School of Nursing
University of British Columbia
T201-2211 Wesbrook Mall
Vancouver, BC, V6T 2B5
Canada
Phone: 1 604 822 7417
Email: lori.block@ubc.ca

Abstract

Background: The World Health Organization is in the process of developing an international administrative classification for health called the International Classification of Health Interventions (ICHI). The purpose of ICHI is to provide a tool for supporting intervention reporting and analysis at a global level for policy development and beyond. Nurses represent the largest resource carrying out clinical interventions in any health system. With the shift in nursing care from hospital to community settings in many countries, it is important to ensure that community nursing interventions are present in any international health information system. Thus, an investigation into the extent to which community nursing interventions were covered in ICHI was needed.

Objective: The objectives of this study were to examine the extent to which International Classification for Nursing Practice (ICNP) community nursing interventions were represented in the ICHI administrative classification system, to identify themes related to gaps in coverage, and to support continued advancements in understanding the complexities of knowledge representation in standardized clinical terminologies and classifications.

Methods: This descriptive study used a content mapping approach in 2 phases in 2018. A total of 187 nursing intervention codes were extracted from the ICNP Community Nursing Catalogue and mapped to ICHI. In phase 1, 2 coders completed independent mapping activities. In phase 2, the 2 coders compared each list and discussed concept matches until consensus on ICNP-ICHI match and on mapping relationship was reached.

Results: The initial percentage agreement between the 2 coders was 47% (n=88), but reached 100% with consensus processes. After consensus was reached, 151 (81%) of the community nursing interventions resulted in an ICHI match. A total of 36 (19%) of community nursing interventions had no match to ICHI content. A total of 100 (53%) community nursing interventions resulted in a broader ICHI code, 9 (5%) resulted in a narrower ICHI code, and 42 (23%) were considered equivalent. ICNP concepts that were not represented in ICHI were thematically grouped into the categories family and caregivers, death and dying, and case management.

Conclusions: Overall, the content mapping yielded similar results to other content mapping studies in nursing. However, it also found areas of missing concept coverage, difficulties with interterminology mapping, and further need to develop mapping methods.

(J Med Internet Res 2019;21(6):e12847) doi:10.2196/12847

KEYWORDS

World Health Organization; classification; nursing informatics; medical informatics; data collection; terminology; community health services; standardized nursing terminology

XSL•FO
RenderX

Introduction

Introduction and Objectives

The digitalization of health care information is increasing rapidly. The use of standardized terminologies and classifications to unambiguously represent this information is a fundamental principle in the field of clinical and biomedical informatics [1]. The World Health Organization Family of International Classifications (WHO-FIC) contains a suite of standardized administrative classification products, which are used internationally and nationally to statistically report on the health and well-being of individuals, families, communities, and populations [2]. The WHO-FIC includes the International Classification of Diseases (ICD), the International Classification of Functioning, Disability and Health (ICF), and the International Classification of Health Interventions (ICHI; in development) [2].

ICHI is the newest classification of this group, and its purpose is to provide a common tool for reporting and analyzing health care interventions [3]. A series of international evaluative projects had been planned for the beta-1 release (eg, terminology mapping and standard case reporting) [4]. The goal of the evaluation projects was to ensure the terminology is (1) robust enough to capture interventions provided across the continuum, (2) appropriate to cover interventions provided by different health care disciplines, (3) has a functional browser tool, and (4) has the depth of educational and training material sufficient to support its future use [4]. Evaluations and releases of the ICHI beta version are ongoing, with a future goal of seeking World Health Assembly approval in 2019 [5].

This descriptive study represents an international evaluative project. Its objectives were to (1) examine the ability of ICHI to represent community nursing interventions found in the International Classification for Nursing Practice (ICNP), (2) provide recommendations for content development, and (3) support continued advancements in understanding the complexities of knowledge representation in standardized clinical terminologies and classifications. In this context, a community nursing intervention refers to the actions carried out by nurses practicing in a community setting to support the health and well-being of patients, families, communities, or populations [6-8]. The multiple research methods used to achieve these research objectives were based on a content mapping approach. Specifically, 2 clinical experts individually matched equivalent (or near equivalent) concepts from ICNP to ICHI. The results were compared and reviewed until matching consensus was reached between the 2 coders. This study is unique in that it is the first to bring a community nursing care perspective to the evaluation of ICHI, informing broader discussions about the representation of health care activities and resourcing in administrative classifications. To the best of our knowledge, it is also the first published study to evaluate aspects of the 2017 ICHI beta-1 release.

Background

Community Nursing

With rapid population growth occurring worldwide, health care systems are challenged both socially and economically, with changing demographics, shifting disease patterns, increased prevalence of chronic diseases, and financial reforms [9]. The delivery of health care services outside of acute care centers is necessary to manage these complex phenomena. Therefore, community nursing is an essential global service. The WHO defines community nursing as a service which “combines the skills of nursing, public health and some phases of social assistance and functions as part of the total public health program for the promotion of health, the improvement of the conditions in the social and physical environment, rehabilitation of illness and disability” [10,11].

Nurses practicing in the community context provide care that directly improves the health outcomes of individuals, families, communities, and populations [12]. This can be attributed to the ethos of community nursing, where work is founded on the principles of social justice, holistic care, equity, ethics, community capacity building and empowerment, and action upon the intersectoral determinants of health [12]. The types of interventions community nurses provide are extensive and can include home visits for new baby and family care, school classes on the topic of sexual health, wound care, interventions that address preventing elder abuse, and advocacy for health and wellness initiatives [13]. Despite the increasing international recognition and support for this nursing service, there remains a limited understanding of the full impact of community nursing on health outcomes [14-16].

The International Classification of Health Interventions

Since its early initiation, ICHI was envisioned as a standardized classification system to describe health care interventions provided by health professionals [17]. To structure the context of this work, developers defined *health intervention* to mean “an act performed for, with or on behalf of a person or a population whose purpose is to assess, improve, maintain, promote or modify health, functioning or health conditions” [18]. The purpose of ICHI was to facilitate the comparison of semantically equivalent information at local, national, or international levels; act as a national classification for countries where no existing (or outdated) intervention classification systems existed; and complement the existing WHO-FIC classifications, ICD and ICF [17,18].

In 2007, working groups within the WHO-FIC began to direct the development of this international classification. A categorial structure, developed by the European Standard Body CEN TC 251/International Standards Organization TC 215 group, was used to build and define the included ICHI content including a framework that defined the way concepts would be related to each other [4,17-19].

Semantic categories within ICHI are structured into 3 axes:

XSL•FO
RenderX

- Target: the semantic categories that the intervention (action) is carried out on, to, or with (eg, person, family, and community)
- Action: the semantic categories describing the intervention done by the actor to the target (eg, assessment, treating, assisting, and informing)
- Means: the semantic categories defining the intervention (action) method or process (eg, method, approach, and technique)

In 2012, an alpha version of the classification became available (in Excel format) to affiliated researchers and partners [18]. After several years, iterations, and evaluative projects, the beta version of ICHI became available to the public through a functional Web browser. This browser allowed users to search through over 7000 concepts in 4 category sections [3,4].

1. Interventions on Body Systems and Functions (eg, biomedical body systems)
2. Interventions on Activities and Participation Domains (eg, activities of daily living)
3. Interventions on the Environment (eg, products, services, and systems)
4. Interventions on Health-related Behaviours (eg, safety and lifestyle)

In a recent release of ICHI, developers defined the use of extension codes allowing for the broadening of the intervention classification (eg, assistive and therapeutic products) [4]. This inclusion has allowed for the classification to grow and to continue in relevance [20]. In late 2018, ICHI released a beta-2 version, which included a noted increase in concept coverage and updated resource materials.

The International Classification for Nursing Practice

The International Council of Nurses (ICN) represents around 20 million nurses in more than 130 nursing associations across the world [21]. ICN develops and distributes ICNP, a standardized terminology system for nursing [22,23]. ICNP conforms to 18104:2014 Health informatics—Categorial structures for representation of nursing diagnoses and nursing actions in terminological systems (previously published as ISO 18104:2003) [24,25]. As a formal standardized nursing terminology, ICNP provides a polyhierarchical framework into which nursing diagnoses, interventions, and outcomes are structured and coded for multiple uses [26].

Since 2005, ICNP has utilized the Web Ontology Language to permit automated description logic reasoning, ensure coherence, and support the classification development [27]. Due to its robustness and compliance to international standards, ICNP is widely recognized as a standard terminology appropriately suited to describe the professional practice of nursing. The WHO has included ICNP as a related classification in the WHO-FIC, using it to extend coverage into the domain of nursing [28].

As an invested partner in the advancement of ICHI, ICN has maintained a working relationship with the ICHI development task force. For example, in 2016, researchers mapped 100 frequently recorded ICNP nursing interventions from acute care settings to the 2015 ICHI alpha release [29]. The purpose was to evaluate the degree of ICNP content coverage in ICHI as well as provide recommendations for additions and changes. The researchers in that study found that 80% of ICNP concepts were represented in ICHI. They also found missing content coverage, ambiguities in concept description, and uncertainties in the semantic matching [30].

Methods

Research Design

This is a descriptive research study. The presented work was conducted using a content mapping approach (the most common method used to perform terminology mapping [29,31-35]) in 2 main phases in July and August, 2018. In phase 1, 2 coders completed independent content mapping activities. In phase 2, the 2 coders compared each list and discussed content matches until consensus on ICNP-ICHI match and on mapping relationship was reached. Additional details about these phases are included below.

The community nursing interventions used in this study were derived from the ICNP Community Nursing Catalogue. This catalogue was developed in 2011, updated most recently in 2017, and created in partnership between the Scottish Government and the ICN [36]. The ICN Guidelines for Catalogue Development encourages worldwide validation through global use. The ICNP Community Nursing Catalogue contains 187 community nursing interventions [36]. These interventions (source) were used to identify if there were any equivalent ICHI precoordinated interventions (target) in the draft 2017 beta-1 release.

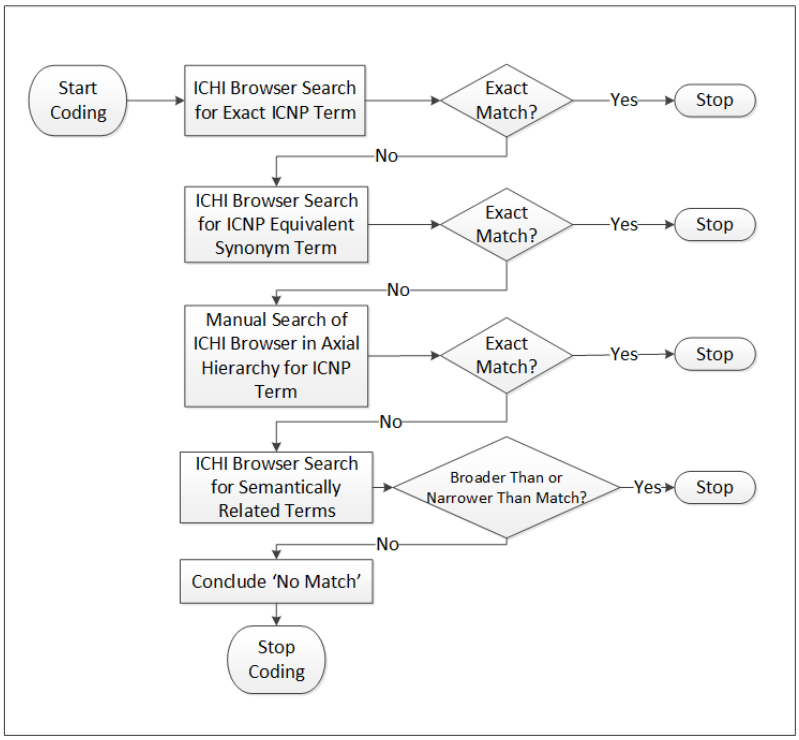
This study did not require research ethics board review as it had no human subjects or materials and was considered a quality assurance and quality improvement evaluation [37,38].

Phase 1: Independent Content Mapping

In phase 1, 2 coders (LB, GS) independently mapped 187 ICNP community nursing interventions to ICHI. The mapping process used by each coder to identify a possible ICNP match to an ICHI intervention was completed using the ICHI online browser and followed the method outlined in Figure 1. For example, if exact or equivalent terms were not immediately found in the ICHI browser search bar, the coders manually searched through the axial categories (eg, interventions on body systems and functions), drilling down through the hierarchical layers (eg, interventions on the integumentary system) until a match (or not) was found. These mapping processes facilitated different mechanisms to manage the search of concepts among the thousands of concepts available to view in the ICHI browser. Different mapping relationships were further considered as exact or equivalent (eg, dog-dog), broader than (eg, dog-mammal), or narrower than (eg, dog-Siberian husky) based on their semantic representation.

XSL•FO
RenderX

Figure 1. Decision process for mapping International Classification for Nursing Practice (ICNP) to International Classification of Health Interventions (ICHI).



The coding was performed in batches to ensure consistency in process and to allow the coders to refine the process over time. This was a mechanism that was established to improve the quality and reliability of the mapping process overall. In the first batch, a systematic sampling method was used to mark every 20th ICNP intervention for a total of 10 (n=10) ICNP intervention codes. This small number allowed the coders to refine the mapping process without having a potentially negative influence on the level of agreement calculated at the end of the study. In the second batch, a total of 30 (n=30) different interventions were selected for coding. This number was selected as it allowed for an additional opportunity to include more types of interventions for refinement in the mapping process. Finally, the remaining interventions (n=147) were coded in the final batch. Other members of the team (LC, NH) were regularly consulted throughout this mapping process and acted to ensure the decision process (Figure 1) was maintained. The mapping took place over a period of 2 months (July and August of 2018).

Phase 2: Reaching Consensus

In phase 2, the independent mapping results were compiled into 1 shared spread sheet. The file contained a list of all ICNP interventions from a particular batch and the matched ICHI intervention from each coder. A percentage agreement between the 2 coders was calculated for each batch. When the coders had different findings from one another, a discussion was carried out until agreement of 1 mapping match was met. The coders also collectively determined the type of mapping relationship for each concept match (equivalent, broader than, narrower than, or no match). These methods are typically followed in content mapping methods to resolve disagreements and come to consensus [29]. As a result, a single ICHI intervention (or no match) was identified for each ICNP intervention. Once completed, final mapping results were presented and discussed among the entire research team, providing opportunity to examine themes and trends of the findings.

Results

Phase 1: Independent Content Mapping

In phase 1, independent coding was completed for all of the ICNP interventions. The percentage agreement between the 2 coders was 47% (n=88). There was no agreement between the coders in the remaining cases (n=99). Table 1 shows examples of cases where the coders identified the same ICHI code, where the coders both identified no ICHI code and where there was no initial mapping agreement.

XSL•FO
RenderX

Table 1. Phase 1 examples of independent content mapping results.

International Classification for Nursing Practice source term/code	ICHI[a] term/code by coder #1	ICHI term/code by coder #2	Coding result
10030440 Advising About Employment	SU2.PN.ZZ Advising about work and employment	SU2.PN.ZZ Advising about work and employment	Agreement (map)
10024570 Supporting Caregiver	No ICHI match identified	No ICHI match identified	Agreement (no map)
10031062 Counselling Patient	PZB.PP.ZZ counselling, not elsewhere classified	No ICHI match identified	Disagreement (different ICHI code was identified)

[a]ICHI: International Classification of Health Interventions.

Phase 2: Reaching Consensus

During phase 2, consensus was achieved for all ICNP interventions (source) through discussion between the 2 coders. A total of 151 cases (81%) of ICNP intervention concepts resulted in an ICHI match. A total of 36 cases (19%) of ICNP intervention concepts resulted in no ICHI match. In the cases where an ICHI match was identified, a conversation ensued about whether ICHI was equivalent to ICNP, whether ICHI was narrower than ICNP, or whether ICHI was broader than ICNP. A summary of the findings and examples are shown in Table 2. Within content mapping methodology, this is a typical approach to identifying equivalency [29,32-35,39].

After the 2 coders completed their mapping consensus work, results were shared with the full research team. As a group, we examined missing ICNP concepts and found thematic groupings that are important practice areas for community nursing. These include intervention concepts related to family and caregivers, death and dying, and case management (Multimedia Appendix 1).

Table 2. Summary of mapping results in phase 2 and examples of mapping specificity.

Mapping result	Statistics, N (%)	Example	
		Source: ICNP[a] intervention term	Target: ICHI[b] code and term
ICHI was equivalent to ICNP	42 (23)	10030558 Assessing bowel continence	KTK.AA.ZZ Assessment of defecation functions
ICHI was narrower than ICNP	9 (5)	10032994 Teaching about effective parenting	SSK.PM.ZZ Education about parent-child relationships
ICHI was broader than ICNP	100 (53)	10030429 Administering vaccine	DTB.DB.AE Other immunization, not elsewhere classified
No match	36 (19)	10032859 Supporting family coping process	(none found)

[a]ICNP: International Classification for Nursing Practice.

[b]ICHI: International Classification of Health Interventions.

Discussion

Principal Findings

The inclusion of community nursing interventions in administrative classifications is essential when evaluating the health and well-being of individuals, families, communities, and populations. The results of this study indicated that 151 of 187 (81%) ICNP community nursing intervention concepts were represented (equivalent, broader, and narrower matches combined) in the beta-1 release of ICHI. Although there is no industry gold standard with which to judge these results, we suggest the representation of community nursing interventions in ICHI appears encouraging. For the 36 (19%) concepts that did not have matches in ICHI, further analysis revealed (1) instances where ICHI was missing representative concepts and (2) inherent differences in terminology system design [29,30]. In addition, the results highlight key considerations related to the representation of knowledge in administrative terminology systems.

Missing Concept Coverage in International Classification of Health Interventions

A total of 36 ICNP intervention concepts were not represented in the ICHI classification. After examining these missing concepts in greater detail, we were able to thematically group several of the intervention concepts into “family and caregivers,” “death and dying,” and “case management.” Inclusion of concepts in ICHI, which consider these themes, is recommended to ensure related concepts are available for administrative reporting and analysis. A focus on the collection of relevant information about community health care provision is necessary to gain knowledge about general health service provision [9].

It is within the scope of practice for community nurses to care for the families and caregivers of a patient [40-45]. In our sample of 187 community nursing interventions, 10 ICNP concepts related to family or caregivers were not represented in ICHI (ie, *10032859 Supporting Family Coping Process; 10032068 Monitoring For Impaired Family Coping*). In particular, this was noted for those concepts specific to community nursing interventions for caregivers of young children (ie, *10032837 Supporting Caregiver During Weaning;*

XSL•FO
RenderX

10033093 Teaching Caregiver About Toilet Training 10032973 Teaching Infant Massage). This practice is often performed by visiting nurses concerned about the functioning and development of young families. Mapping difficulties were also noted when attempting to match ICNP concepts with the specific word "caregiver," as ICHI uses different terms in target descriptions (eg, family, friend, peers, colleagues, neighbors, and community members). Although each of these ICHI target terms could be a "caregiver," in practice, they are not always equivalent. Caring for the caregiver and family is essential to the overall health of a population and necessary to account for in administrative classifications [40-45].

Another area with missing content coverage was noted for those specific ICNP intervention concepts on "death" and "dying" (ie, *10041254 Supporting Dignified Dying; 10033296 Verifying Death*). In the ICHI beta-1 version, no codes specifically used these terms, or even the broader terms of "palliative care," "hospice," or "end of life." This area of practice has always been part of nursing and is increasingly viewed an essential service in the community setting [13]. Cultural, legal, and practice changes are also occurring on this topic of end of life care. For example, in Canada, medical assistance in dying is a legally administered intervention provided by physicians and nurse practitioners and is supported by other health care providers such as registered nurses [46]. Ensuring the representation of appropriate end of life concepts in administrative classifications is necessary as it supports the evaluation of health interventions provided in the community setting.

A theme emerged related to missing content for community nursing "case management". Case management is the coordination of a wide variety of services, which benefit the care of individuals, families, and communities [13]. For example, the role of community nursing in case management activities may include screening of health and functional needs, arranging services, planning care, ongoing reassessment, and provision of continuity between services [13]. In the report *Crossing the Quality Chasm* [47], the need to improve the organization and coordination of care around the needs of a person was stated as a measure to improve the health care system. Though the mapping between ICNP and ICHI did find matches between related concepts (ie, *10030455 Advising About Housing*), several were not found (ie, *10032598 Referring To Housing Service; 10030625 Assessing Housing Condition; 10030493 Arranging Transport of Device*). These missing concepts describe the type of ongoing case management community nurses provide on behalf the persons outside of institutionalized care. It is again recommended that case management intervention concepts continue to be developed and added to administrative classification systems as a means to increase our understanding and inform future health care decisions.

Foundational Design Decisions of a Classification System

The foundational design of a classification or terminology system considers scope, hierarchical orientation, concept granularity, and concept placement. Standards such as ISO 18104:2014 Health informatics--Categorial structures for representation of nursing diagnoses and nursing actions in terminological systems, direct design decisions. For example, ICHI concepts are required to include a defined target, action, and means. ICNP interventions are required to have a target and action but no means. When researchers conduct interterminology mapping exercises, discord between concept representations may be related to these foundational development decisions.

In this mapping activity, several missing ICNP matches were related to differences in concept granularity (ie, specificity or level of detail for related concept). For example, the ICNP concept *10033126 Teaching Patient* was determined to have "no match" in ICHI. This was not because of the lack of codes in ICHI, which could be used to describe patient education. Rather, the ICNP concept was "broader than" what was available in ICHI. One may then ask, why not choose an ICHI concept that was more specific and call it a "narrower match?" The ICNP concept *10033126 Teaching Patient* could have been a "narrower match" to over 300-specific ICHI educational concepts. Practically speaking, the terminology coders could not make a meaningful one-to-one match. The following examples represent additional "broader than" ICNP concepts that did not have meaningful matches in ICHI.

- 10030673 Assessing During Encounter
- 10024570 Supporting Caregiver
- 10032844 Supporting Family
- 10031912 Managing Disease
- 10031965 Managing Symptom
- 10033086 Teaching Caregiver
- 10033126 Teaching Patient

This example highlights the complexities of knowledge representation when attempting to map terminologies of varying granularity and overlapping coverage. When decisions are made on how a terminology or classification is to be foundationally structured, and then mapped to another with a different foundational base, clashes in semantic matching may be part of the expected results.

Representation of Community Nursing Practice

As noted above, a total of 151 (81%) ICNP community nursing interventions are represented in ICHI. Two-thirds of these concept matches were classified as "broader than" (ie, meaning that an ICNP concept could fit as a "child" into the broader ICHI "parent" concept). From the vantage of developing an administrative classification to represent health, it can be understood that there has to be a threshold of low specificity to allow for a higher aggregation of data. However, the question remains as to whether these "broader than" ICHI concepts satisfactorily represent nursing care interventions and at what point knowledge representation turns from meaningful coverage to diluted meaninglessness.

XSL•FO
RenderX

Table 3. International Classification for Nursing Practice concepts not elsewhere classified.

International Classification for Nursing Practice concept	"Broader than" International Classification of Health Interventions concept
10031117 Diabetic Ulcer Care	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10031690 Malignant Wound Care	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10032420 Pressure Ulcer Care	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10032863 Surgical Wound Care	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10033208 Traumatic Wound Care	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10033254 Ulcer Care	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10030710 Assessing Risk For Pressure Ulcer	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10030723 Assessing Risk For Transfer Injury	LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified
10031931 Managing Postpartum Care	NUE.ZY.ZZ Other interventions on functions related to pregnancy, not elsewhere classified
10031949 Managing Prenatal Care	NUE.ZY.ZZ Other interventions on functions related to pregnancy, not elsewhere classified
10030706 Assessing Risk For Depressed Mood During Postpartum Period	NUE.ZY.ZZ Other interventions on functions related to pregnancy, not elsewhere classified
10031769 Managing Postpartum Depressed Mood	NUE.ZY.ZZ Other interventions on functions related to pregnancy, not elsewhere classified
10031805 Managing Enuresis	NTD.ZY.ZZ Other interventions on urination function, not elsewhere classified
10031879 Managing Urinary Incontinence	NTD.ZY.ZZ Other interventions on urination function, not elsewhere classified
10033135 Teaching Self-Catheterisation	NTD.ZY.ZZ Other interventions on urination function, not elsewhere classified
10033277 Urinary Catheter Care	NTD.ZY.ZZ Other interventions on urination function, not elsewhere classified
10035958 Facilitating Grief	AUD.ZY.ZZ Other interventions on emotional functions, not elsewhere classified
10031711 Managing Anxiety	AUD.ZY.ZZ Other interventions on emotional functions, not elsewhere classified
10031851 Managing Negative Emotion	AUD.ZY.ZZ Other interventions on emotional functions, not elsewhere classified

In the case of community nursing skin and wound care concepts, 90% were matched to ICHI as "broader than" (10% no matches). For example, 8 skin and wound care ICNP concepts were rolled up into the closest ICHI match, *LZZ.ZY.ZZ Other interventions on integumentary system, not elsewhere classified*. Similar outcomes were found for ICNP concepts related to prenatal and postpartum care, continence and catheter care, and supporting care for grief and anxiety (Table 3). If these concepts were subsequently mapped against health care data, the knowledge represented would be so vague that extracting knowledge back out of it could be lost. These are important considerations, especially as these concepts not only represent the care provided by community nursing but also many other health care professional groups. ICHI is being developed for countries to report and analyze on health interventions [3]. It is recommended, therefore, that ongoing work continues to evaluate the practical use (eg, to support resourcing) of those concept groups frequently mapped as "broader than," to ensure the meaningful representation of health care phenomena is available in administrative classifications [20].

Mapping Methods of Coding and Consensus

There is no agreed upon method of mapping concepts from a source to a target classification or terminology. Multiple examples of mapping clinical content between interterminology groups, datasets to terminologies, or raw clinical content exist [23,29,39,48,49]. In this study, we presented a method of using 2 coders to manually map 187 concepts from 1 standardized

XSL•FO
RenderX

clinical terminology to another standardized clinical classification. This mapping exercise was greatly aided by both the ICHI and ICNP publicly available Web browsers.

During phase 1, only 47% (n=88) of the concept matches were the same between the 2 coders; this increased to 100% in phase 2. Although the percentage agreement was low at the beginning, statistically suggesting weakness in the initial findings [50], the science of clinical informatics is still maturing and has yet to demonstrate how this value fully impacts the reliability of mapping results [51]. It is possible that this lower agreement was related to large number of target concepts (eg, ICHI beta-1 version had 7000 concepts), differences in concept understanding (eg, differences between counseling, advising, education, and emotional support), and different levels of experience in mapping ICNP and ICHI content.

To increase the trustworthiness of the content mapping process, batches of coding and consensus gathering were completed to provide a quality assurance mechanism by allowing the coders to further clarify and consistently manage the coding process. During the first batch of intervention discussions, senior researchers in field (LC, NH) provided coaching regarding how to consistently manage the coding process. This acted as a quality control mechanism before the remainder of the content mapping was completed. The remaining batches were discussed and resolved without the senior researchers' presence. The browser tool was also used throughout the consensus discussions between the 2 coders. In particular, when debating between 2 different ICHI concepts, the coders would consult the concept definition and inclusion fields found when clicking the ICHI concept. This discussion process facilitated a final 100% agreement of mapping results by the end of phase 2.

Finally, it should be noted that the coders are registered nurses with both clinical practice and content mapping experience. This facilitated the coders to use explicit knowledge to understand concept meaning in context to community care, to find concept synonyms (eg, step 2 in the mapping method process), and to easily navigate the ICHI Web browser. It is outside the scope of this study to examine how tacit knowledge, experiential judgment, or social relationships (eg, consensus agreement) may have contributed to the coders' mapping choices. Future researchers may wish to examine the influence of these variables on concept terminology mapping results. For example, those research methods that capture the decision-making process of a coding task (eg, Think Aloud protocols) may potentially be a fruitful line of inquiry.

Limitations

There are limitations related to the repeatability of this study. Though we have attempted to be clear and robust in the methods and processes used to map the ICNP community content to ICHI, the findings may have been different had there been different coders or different versions of classifications. For example, the ICHI beta-1 version (utilized over the coding period of Summer 2018) was updated in October 2018 to ICHI beta-2, increasing clinical concepts from approximately 7000 to 8000. It is possible that the rate of agreement between the 2 classifications would be different with updated and ongoing versions.

Conclusions

The collection of standardized information from electronic health records is used to help institutions to determine priorities and effective allocation of resources [10]. As the shift toward preventative and community-based health care increases, so too does the need for health organizations to have well informed administrative data about this domain. The work presented in this study helps advance the representation of community nursing concepts in administrative datasets, a relatively new challenge for nursing informatics; however, although this is a necessary step, it does not guarantee that these data will be utilized in reporting. Continued work is necessary to champion and value the work of community nursing, which will further contribute to a wholesome account of the health and well-being of individual, families, communities, and populations.

Acknowledgments

The authors would like to thank the ICHI task force, chaired by Richard Madden and Lyn Hanmer, for the opportunity to contribute to the evaluation of community nursing representation in the beta version of ICHI.

Conflicts of Interest

NRH has acted as a consultant to the International Council of Nurses.

Multimedia Appendix 1

International Classification for Nursing Practice to International Classification of Health Interventions community nursing mapping results.

[PDF File (Adobe PDF File), 180KB - jmir_v21i6e12847_app1.pdf]

References

1. Sensemeier J. Health data standards: development, harmonization, interoperability. In: Saba V, McCormick K, editors. Essentials of Nursing Informatics. Sixth Edition. New York, New York, USA: McGraw Hill; 2015.

XSL•FO
RenderX

2. Madden R, Sykes C, Ustun TC. World Health Organization. 2012. World health organization family international classifications: Definition, scope and purpose URL: https://www.who.int/classifications/en/FamilyDocument2007.pdf?ua=1 [accessed 2019-05-21] [WebCite Cache ID 78XEPHExE]
3. Donada M, Della Mea V, Cumerlato M, Rankin N, Madden R. A system for supporting development and update of the international classification of health interventions (ICHI). Stud Health Technol Inform 2018;247:895-899. [Medline: 29678090]
4. Fortune N, Madden R, Almborg A. Use of a new international classification of health interventions for capturing information on health interventions relevant to people with disabilities. Int J Environ Res Public Health 2018 Dec 17;15(1):145 [FREE Full text] [doi: 10.3390/ijerph15010145] [Medline: 29342077]
5. Rodrigues JM, Kim S, Paviot BT, Lee JJ, Aljunid S. How to link SNOMED CT procedure and WHO international classification of health interventions (ICHI). Stud Health Technol Inform 2017;236:40-47. [Medline: 28508777]
6. Deal LW. The effectiveness of community health nursing interventions: a literature review. Public Health Nurs 1994 Oct;11(5):315-323. [Medline: 7971696]
7. Snyder M, Egan EC, Nojima Y. Defining nursing interventions. Image J Nurs Sch 1996;28(2):137-141. [Medline: 8690430]
8. International Council of Nurses. International Council of Nurses. 2018. Why use ICNP? URL: https://www.icn.ch/sites/default/files/inline-files/Benefits [accessed 2019-05-21] [WebCite Cache ID 78XH6Sdcm]
9. World Health Organization Regional Office for South-East Asia.. World Health Organization. 2010. A framework for community health nursing education URL: http://www.searo.who.int/entity/nursing_midwifery/documents/SEA-NUR-467/en/ [accessed 2019-05-21] [WebCite Cache ID 78XLWTQJj]
10. WHO Expert Committee on Community Health Nursing & World Health Organization. World Health Organization. 1974. Community health nursing : report of a WHO expert committee [meeting held in Geneva from 30 July to 5 August 1974] URL: https://apps.who.int/iris/handle/10665/41121 [accessed 2019-05-25] [WebCite Cache ID 78eI3ENno]
11. World Health Organization. Enhancing the role of community health nursing for universal health coverage URL: https://www.who.int/hrh/resources/health-observer18/en/ [accessed 2019-05-21] [WebCite Cache ID 78XOaX3Dk]
12. Wolf K, Breakey S, Corless IB, Meedzan N, Nicholas PK. The intersection of global health and community/public health nursing. In: Global Health Nursing in the 21st Century. New York, New York: Springer Publishing Company; 2015.
13. Nies M. Promoting the health of populations. In: Nies M, McEwen M, editors. Community/Public Health Nursing, 7th Edition. Community/Public Health Nursing - E-Book: Elsevier Health Sciences. Kindle Edition; 2019.
14. Barret A, Terry D, Le Q, Hoang H. Factors influencing community nursing roles and health service provision in rural areas: a review of literature. Contemp Nurse 2016;52(1):119-135. [Medline: 27264878]
15. Atherton I, Lynch A, Williams AJ, Witham MD. Barriers and solutions to linking and using health and social care data in Scotland. Br J Soc Work 2015 Jun 2;45(5):1614-1622. [doi: 10.1093/bjsw/bcv047]
16. Nagle L, White P. CIHI. Proceedings from the 2018 national nursing data standards symposium URL: https://www.cihi.ca/sites/default/files/document/nnds_2018_proceedings.pdf [accessed 2019-05-21] [WebCite Cache ID 78XP1WKsD]
17. Paviot B, Madden R, Moskal L, Zaiss A, Bousquet C, Kumar A, et al. Development of a new international classification of health interventions based on an ontology framework. Stud Health Technol Inform 2011;169:749-753. [Medline: 21893847]
18. World Health Organization. 2018. ICHI Browser Beta Version Internet URL: https://mitel.dimi.uniud.it/ichi/ [accessed 2019-05-21] [WebCite Cache ID 78XPFFSss]
19. Rodrigues J, Kumar A, Bousquet C, Trombert B. Standards and biomedical terminologies: the CEN TC 251 and ISO TC 215 categorial structures: a step towards increased interoperability. Stud Heal Technol Informatics 2008;136:857. [Medline: 18487839]
20. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998 Nov;37(4-5):394-403 [FREE Full text] [Medline: 9865037]
21. International Council of Nurses. Who We Are. Internet URL: https://www.icn.ch/who-we-are [accessed 2019-05-22] [WebCite Cache ID 78ZOoyI3u]
22. International Council of Nurses. International Classification for Nursing Practice (ICNP) Catalogue. Geneva: International Council of Nurses; 2015:1-16.
23. Kim T. Standards. In Health Informatics: An Interprofessional Approach. St Louis: Elsevier Mosby; 2014.
24. International Council of Nurses. 2010. International Organisation for Standards (ISO) Internet URL: https://www.icn.ch/what-we-do/projects/ehealth-icnp [accessed 2019-05-22] [WebCite Cache ID 78ZPsaM43]
25. Kobs A. ISO 9001: a support for professional nursing practice. Nurse Lead 2006;4(4):2000. [doi: 10.1016/j.mnl.2006.05.007]
26. Marin HD. Categorical structure analysis of ISO 18104 standard in nursing documentation. Acta Paul Enferm 2013;26(3):299-306. [doi: 10.1590/S0103-21002013000300016]
27. Hardiker N, Coenen A. Interpretation of an international terminology standard in the development of a logic-based compositional terminology. Int J Med Inform 2007;76:274. [Medline: 17600764]
28. World Health Organization. 2017. International Classification for Nursing Practice (ICNP) URL: https://www.who.int/classifications/icd/adaptations/icnp/en/ [accessed 2019-05-24] [WebCite Cache ID 78c0ITUqt]

XSL•FO
RenderX

29. Fortune N, Hardiker NR, Strudwick G. Embedding nursing interventions into the World Health Organization's International Classification of Health Interventions (ICHI). J Am Med Inform Assoc 2017 Jul 1;24(4):722-728. [doi: 10.1093/jamia/ocw173] [Medline: 28339684]
30. Ohannessian R, Fortune N, Rodrigues J, Moulin T, Derex L, Madden R. Coding acute stroke care and telestroke with the international classification of health interventions (ICHI). Int J Med Inform 2017;108:9. [Medline: 29132637]
31. Kim TY, Hardiker N, Coenen A. Inter-terminology mapping of nursing problems. J Biomed Inform 2014 Jun;49:213-220 [FREE Full text] [doi: 10.1016/j.jbi.2014.03.001] [Medline: 24632297]
32. Kuo C, Yen M. Cross-mapping ICNP terms with Taiwanese gynecological nursing records. J Nurs Res 2006;14(4):271-278. [Medline: 17345756]
33. Goosen W. Cross-mapping between three terminologies with the international standard nursing reference terminology model. Int J Nurs Terminol Classif 2006;17(4):153-164. [Medline: 17117945]
34. Cubas M, Carvalho C, Malucelli A, Denipote A. Cross-mapping of terms from the action axis between different nursing classifications. Rev Bras Enferm 2011;64(2):248-253. [Medline: 21755206]
35. Hyun SP. Cross-mapping the ICNP with NANDA, HHCC, Omaha system and NIC for unified nursing language system development. Int Nurs Rev 2002;49(2):99-110. [Medline: 12094837]
36. International Council of Nurses. 2017. ICNP Community Nursing Catalogue URL: https://www.icn.ch/what-we-do/projects/ehealth-icnp/about-icnp/icnp-catalogues [accessed 2019-05-26] [WebCite Cache ID 78eK7Wauj]
37. The Interagency Advisory Panel on Research Ethics (PRE). 2017. TCPS 2 (2014)— the latest edition of Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans. Humans Internet URL: http://www.pre.ethics.gc.ca/eng/policy-politique/initiatives/tcps2-eptc2/Default/ [accessed 2019-05-24] [WebCite Cache ID 78c0qJYCo]
38. NHS Health Research Authority. 2017. UK Policy Framework for Health and Social Care Research URL: https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/uk-policy-framework-health-social-care-research/ [accessed 2019-05-24] [WebCite Cache ID 78c1AkpfI]
39. Kim TY, Hardiker N, Coenen A. Inter-terminology mapping of nursing problems. J Biomed Inform 2014 Jun;49:213-220 [FREE Full text] [doi: 10.1016/j.jbi.2014.03.001] [Medline: 24632297]
40. Kim E, Yeom H. Influence of home care services on caregivers' burden and satisfaction. J Clin Nurs 2016;25:1683-1692. [Medline: 26898819]
41. Moorman S, Macdonald C. Medically complex home care and caregiver strain. Gerontologist 2013;53(3):407-417. [Medline: 22563000]
42. Berthelsen C, Kristensson J. The content, dissemination and effects of case management interventions for informal caregivers of older adults: a systematic review. Int J Nurs Stud 2015;52(5):988-1002. [Medline: 25661314]
43. Sprung S, Laing M. Young carer awareness, identification and referral. Br J Community Nurs 2017;22(8):398-406. [Medline: 28767308]
44. Kemp L, Harris E, McMahon C, Matthey S, Vimpani G, Anderson T, et al. Child and family outcomes of a long-term nurse home visitation programme: a randomised controlled trial. Arch Dis Child 2011 Jun;96(6):533-540. [doi: 10.1136/adc.2010.196279] [Medline: 21429975]
45. Sawyer M, Frost L, Bowering K, Lynch J. Effectiveness of nurse home-visiting for disadvantaged families: Results of a natural experiment. BMJ Open 2013;3(4):e002720. [Medline: 23619089]
46. Government of Canada. 2018. Medical assistance in dying URL: https://www.canada.ca/en/health-canada/services/medical-assistance-dying.html [accessed 2019-05-24] [WebCite Cache ID 78c2vYzI0]
47. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
48. Harris M, Langford LH, Miller H, Hook M, Dykes P, Matney SA. Harmonizing and extending standards from a domain-specific and bottom-up approach: an example from development through use in clinical applications. J Am Med Informatics Assoc 2015;22(3):545-552. [Medline: 25670750]
49. Richesson RL, Andrews JE, Krischer JP. Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. J Am Med Inform Assoc 2006;13(5):536-546 [FREE Full text] [doi: 10.1197/jamia.M2093] [Medline: 16799121]
50. Lavrakas P. Encyclopedia of Survey Research Methods. Thousand Oaks: Califage Publications; 2008.
51. Amith M, He Z, Bian J, Lossio-Ventura JA, Tao C. Assessing the practice of biomedical ontology evaluation: gaps and opportunities. J Biomed Inform 2018 Dec;80:1-13 [FREE Full text] [doi: 10.1016/j.jbi.2018.02.010] [Medline: 29462669]

Abbreviations

ICD: International Classification of Diseases
ICF: International Classification of Functioning, Disability and Health
ICHI: International Classification of Health Interventions
ICNP: International Classification for Nursing Practice
WHO-FIC: World Health Organization Family of International Classifications

XSL•FO
RenderX

Edited by G Eysenbach; submitted 16.11.18; peer-reviewed by S Khasnabish, V Bremer; comments to author 14.02.19; revised version received 11.04.19; accepted 02.05.19; published 26.06.19.

Please cite as:
Block LJ, Currie LM, Hardiker NR, Strudwick G
Visibility of Community Nursing Within an Administrative Health Classification System: Evaluation of Content Coverage
J Med Internet Res 2019;21(6):e12847
URL: https://www.jmir.org/2019/6/e12847/
doi:10.2196/12847
PMID:31244480

©Lorraine J Block, Leanne M Currie, Nicholas R Hardiker, Gillian Strudwick. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

When Similarity Beats Expertise—Differential Effects of Patient and Expert Ratings on Physician Choice: Field and Experimental Study

Anne-Madeleine Kranzbühler[1], PhD; Mirella H P Kleijnen[2], PhD; Peeter W J Verlegh[2], PhD; Marije Teerling[3], PhD

[1]Department of Product Innovation Management, Delft University of Technology, Delft, Netherlands

[2]Department of Marketing, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[3]Eneco, Rotterdam, Netherlands

Corresponding Author:
Anne-Madeleine Kranzbühler, PhD
Department of Product Innovation Management
Delft University of Technology
Landbergstraat 15
Delft, 2628 CE
Netherlands
Phone: 31 152783451
Email: a.kranzbuhler@tudelft.nl

Abstract

Background: Increasing numbers of patients consult Web-based rating platforms before making health care decisions. These platforms often provide ratings from other patients, reflecting their subjective experience. However, patients often lack the knowledge to be able to judge the objective quality of health services. To account for this potential bias, many rating platforms complement patient ratings with more objective expert ratings, which can lead to conflicting signals as these different types of evaluations are not always aligned.

Objective: This study aimed to fill the gap on how consumers combine information from 2 different sources—patients or experts—to form opinions and make purchase decisions in a health care context. More specifically, we assessed prospective patients' decision making when considering both types of ratings *simultaneously* on a Web-based rating platform. In addition, we examined how the influence of patient and expert ratings is conditional upon rating volume (ie, the number of patient opinions).

Methods: In a field study, we analyzed a dataset from a Web-based physician rating platform containing clickstream data for more than 5000 US doctors. We complemented this with an experimental lab study consisting of a sample of 112 students from a Dutch university. The average age was 23.1 years, and 60.7% (68/112) of the respondents were female.

Results: The field data illustrated the moderating effect of rating volume. If the patient advice was based on small numbers, prospective patients tended to base their selection of a physician on expert rather than patient advice (profile clicks beta=.14, P<.001; call clicks beta=.28, P=.03). However, when the group of patients substantially grew in size, prospective patients started to rely on patients rather than the expert (profile clicks beta=.23, SE=0.07, P=.004; call clicks beta=.43, SE=0.32, P=.10). The experimental study replicated and validated these findings for conflicting patient versus expert advice in a controlled setting. When patient ratings were aggregated from a high number of opinions, prospective patients' evaluations were affected more strongly by patient than expert advice ($mean_{patient\ positive/expert\ negative}$=3.06, SD=0.94; $mean_{expert\ positive/patient\ negative}$=2.55, SD=0.89; $F_{1,108}$=4.93, P=.03). Conversely, when patient ratings were aggregated from a low volume, participants were affected more strongly by expert compared with patient advice ($mean_{patient\ positive/expert\ negative}$=2.36, SD=0.76; $mean_{expert\ positive/patient\ negative}$=3.01, SD=0.81; $F_{1,108}$=8.42, P=.004). This effect occurred despite the fact that they considered the patients to be less knowledgeable than experts.

Conclusions: When confronted with information from both sources simultaneously, prospective patients are influenced more strongly by other patients. This effect reverses when the patient rating has been aggregated from a (very) small number of individual opinions. This has important implications for how to present health care provider ratings to prospective patients to aid their decision-making process.

XSL•FO
RenderX

(J Med Internet Res 2019;21(6):e12454) doi:10.2196/12454

KEYWORDS

decision making; choice behavior; judgment

Introduction

Background

The rise of the internet, with its abundance of information, has changed human decision making [1,2]. To reduce their uncertainty when facing a choice, people have always sought advice from other people [3-5]. Web-based information sources that provide electronic word of mouth have become ubiquitous, which has also influenced people's health care decisions [6-8]. When choosing a physician, for example, many patients consult health care rating platforms to obtain information about the quality of different providers [9,10]. Such platforms may be provided by governments (eg, United States: medicare.gov, United Kingdom: nhs.uk) or companies (eg, healthgrades.com, betterdoctor.com) and allow people to consult other patients' ratings of institutions, physicians, or specific procedures [11]. Parallel to this increased use of ratings by consumers, these ratings have also gained importance in public health systems; for example, in the United States [12,13]. Medicare social insurance program financially penalizes hospitals that perform poorly in their patient ratings [14,15].

Such uses of patient ratings might be questionable though, because health care, similar to other professional services, is complex and marked by credence qualities [16,17]. Even after it has been performed, it is difficult for patients to assess the actual quality of a health service, especially because laypeople often lack the skills needed to evaluate the complex offerings [18]. Accordingly, when prior patients evaluate a physician on a Web-based rating platform, they tend to highlight easily observable features related to the care, such as a physician's convenient location or friendly office staff [7,19], rather than cure-related outcomes [20]. A positive service experience is integral to health care, in that it contributes to psychological and physiological health [21], but such assessments ignore another essential element: the technical quality of health care procedures [13]. Moreover, though consumers often use their service experience as a proxy for technical quality, these factors are at best related, not substitutive, components of the same assessment [22,23]. To address this problem, many rating platforms include outcome-focused expert judgments based on statistics, such as readmission rates, as complementary sources of information. However, little is known about how consumers combine information from the 2 different sources—namely, similar patients or experts (with detailed knowledge)—to form opinions and make purchase decisions.

This study aims to fill this gap. On the basis of source effects literature, we investigated prospective patients' decision making when simultaneously exposed to other patients' and expert ratings on Web-based platforms. Whereas previous research primarily focused on the effects of *either* user or expert ratings on decision making in various contexts (eg, as seen in some studies [24-27]), we focused on situations where both types of ratings are provided *simultaneously*. This design feature of Web-based rating platforms is not only of increasing practical interest, but it also allows for disagreement between the patient and expert advice, which is of theoretical interest because it forces prospective patients to make a decision about which source they want to follow. The literature on source effects suggests that patients may overly rely on information from similar sources, so that patients may base their decisions more strongly on advice from other patients (who are more similar to them) than from experts, although the latter may be seen as more competent to give advice in the respective situation [27].

We further investigated how the influence of patient and expert ratings is dependent on the number of patient opinions that contributed to the aggregated patient rating (hereafter referred to as "rating volume"). Rating volume has been widely studied in the literature on Web-based reviews and ratings (eg, [11,28-30]), but little is known about how website users make use of this additional cue when simultaneously confronted with different sources of advice. A common design practice of Web-based rating platforms is to present only 1 expert rating, whereas the patient rating typically is based on the aggregated opinion of multiple patients. On the basis of social influence theory, we argue that the influence of patient over expert ratings is dependent on the number of underlying patient opinions that contributed to the overall rating. This is not only an important insight for practitioners who have to decide about the design of such a platform, but it also helps to understand the psychological processes underlying patients' preferences for advice from other patients or experts.

Source Effects

To determine the usefulness of a recommendation, a receiver must evaluate its source [31]. A source's perceived expertise and perceived similarity to the receiver offer the best predictors of its impact on decision making [32,33]. Although other patients can be expected to be perceived as more similar than experts, they most likely lack the experts' detailed knowledge.

Similarity

Source similarity describes the extent to which people resemble one another on certain attributes. Demographic similarity refers to attributes such as gender, age, or socioeconomic status; attitudinal similarity instead pertains to values, attitudes, or lifestyles [32]. Similar sources tend to be particularly influential because they share similar needs and preferences with the receiver and thus deliver relevant information [34,35]. Similarity also leads to greater attractiveness, trust, and understanding [36], which facilitates communication [37]. Therefore, information from similar others is more persuasive and has a stronger influence on decision making than information from dissimilar others (eg, [35,38]). However, socially meaningful similarities can create a perceived bond between people, in addition to the fact that they share a similar experience or situation [39]. In the present health care context, the user might realize that other patients have been in the same situation of

XSL•FO
RenderX

needing a certain treatment. Experts instead might be perceived as being less similar as they do not share the same situation and probably approach the assessment of a service in a more abstract and technical manner.

Expertise

Expertise is the source's "ability to perform product-related tasks successfully" [40]. Perceptions of a source's expertise stem from evaluations of its knowledge, experience, or occupation [41]. In a Web-based environment, this evaluation relies on the limited available cues [42], although most consumers accept an expert source presented on a website and do not investigate the basis of its expertise further [43,44]. The influence of expert sources stems from their expansive knowledge base, compared with nonexpert sources [32,45]. They have greater knowledge about various, alternative offerings because of their enduring involvement in a certain product or service category [46]. In turn, experts' opinions appear to be of high quality, and their advice has a strong, positive impact on receivers (eg, [32,47,48]).

Comparing the Impact of Similarity Versus Expertise

Both perceived expertise and perceived similarity enhance the persuasiveness of a source (eg, [49]), but we do not know which source has the stronger influence: the expert or the similar peer. Research so far compares their impact only indirectly, by employing survey methods or confronting participants with either type. An early study found that consumers are more inclined to follow the advice of a similar but inexperienced salesperson compared with a dissimilar but experienced one [34]. In line with this, expert reviews have been found to reduce consumers' willingness to visit a restaurant's website, whereas consumer reviews increase it [50]. Similarly, when looking for a new doctor, most participants in a US study sought advice from similar sources rather than those with medical expertise [51]. Yet, other studies indicate instead that source expertise is a better predictor of influence [32,52]. A recent meta-analysis affirms a greater impact on sales elasticities from a review provided by an expert rather than a consumer source [53]. Thus, there is conflicting evidence from various contexts, which leads to our main research question:

> *What is the impact of expert and patient ratings when simultaneously provided on a Web-based rating platform?*

The Moderating Role of Rating Volume Information

In the present context of a *simultaneous* presentation of patient and expert ratings, website users might look for additional cues, such as volume information, to help them make inferences about the sources [54,55]. Rating platforms very often not only simultaneously provide star ratings from experts and patients but also inform website users about how many individual patient opinions have been aggregated to compute a rating. However, this ancillary information is usually only given for patient ratings, whereas expert ratings are not further qualified by specific volume information (eg, betterdoctor.com, nhs.uk, and independer.nl). Thus, this additional cue might impact the way website users construe the patient rating in relation to the simultaneously provided expert rating. This design feature leads to the theoretically profound question—whether information about rating volume (ie, the number of opinions of other patients that underlie a physician rating) may constitute an important moderator for the influence of such ratings on other patients' service evaluations.

Social influence literature states that the likelihood that an individual imitates a certain behavior increases with the number of other individuals who display this behavior [56,57]. The high number of such individuals increases the perceived benefits of adopting a certain behavior compared with its perceived costs [58,59]. Numerical information has been found to be automatically encoded by the human brain and thus easy to process relatively independent of individual differences [60]. The numerical dominance of a majority opinion signals its correctness [61,62]. Humans have been found to be especially prone to the influence from a large number of others when it comes to making purchase decisions (eg, [63-65]). Humans engage in such imitative behavior because a high volume increases the diagnosticity of a message's valence and thus its persuasiveness [29]. This might be due to the fact that high consensus signals that those others might possess a piece of information that is not yet available to the decision maker, (ie, "they must all know something I don't") [63,66]. In line with that, the credibility as well as the impact of a group's judgment on an individual have been found to be positively correlated with the group's size—both Web-based and offline (eg, [28,67]). Hence, rating volume is likely to amplify the effect of advice from similar patients (compared with expert advice) on physician evaluation and decision making. In other words: the larger the number of patients who rate a physician, the more influential the (averaged) patient rating. We thus hypothesize the following:

> *The influence of patient (versus expert) ratings on(1) evaluations and (2) usage intentions of a health service is moderated by rating volume, such that with increasing numbers of underlying patient opinions, the influence of the patient over the expert rating increases.*

Methods

Study 1: Field Study

Study 1 relies on clickstream data from an actual US health care rating platform to assess the impact of simultaneously provided patient and expert ratings as well as the moderating role of explicit rating volume information. The US-based rating platform provides expert and patient ratings on more than 1 million doctors, dentists, and eye doctors throughout the United States to help patients find a suitable service provider. For each doctor, an algorithm calculates an expert rating on the basis of the physician's education, experience, training, and referrals by other doctors. It also displays the average Yelp (patient) rating for each doctor. Both rating types are presented next to each other on overview and search results pages from which patients can access the individual doctor profiles. We assessed the impact of both rating types on 2 behavioral actions: the number of times website users viewed a certain doctor's profile and the number

XSL•FO
RenderX

of times they clicked a button to obtain the doctor's contact information.

Data Collection

We obtained clickstream data for 5299 doctors during May 1 to July 31, 2015 from the rating platform's Web analytics tool and aggregated the click-level data (21,897 profile clicks and 1842 call clicks, see below) to the doctor-level (n=5299). We only included doctors whose profiles featured both expert and patient ratings and whose profiles had been viewed at least once by a prospective patient during the data collection period. We extracted how many times each doctor's profile had been viewed and how many times users clicked to obtain the doctor's contact information. Furthermore, for each doctor, we extracted the expert star rating, the Yelp (patient) star rating, and the rating volume; the overall Yelp rating was based on at the moment of the click. In addition, we collected several control variables (see below).

Measures

We have 2 dependent variables: number of profile clicks and number of call clicks, that is, we counted the number of times each doctor profile was viewed by prospective patients in the data collection period. We also determined how many times visitors to a specific doctor's profile clicked on a button to obtain this doctor's phone number. This behavioral measure provides a proxy for doctor choice.

The independent variables reflected expert and patient ratings. Expert ratings were the professional star ratings provided for each doctor, which ranged from 1 to 5 in 0.5 steps (half stars). The average expert rating was very positive (mean 4.27, SD 0.94). Patient ratings were provided in the form of Yelp star ratings on each doctor's profile. The Yelp ratings could change during the data collection period; we therefore used the respective Yelp rating at the moment of each individual click to calculate an average patient rating for each doctor over the data collection period. Patient ratings also varied from 1 to 5, in 0.5 steps, and the average patient rating in our sample was also positive (mean 3.89, SD 1.21). To be able to test the effect of rating volume, we included the number of individual patient opinions the Yelp star rating was based on in our model. This number could also change over the course of the data collection period; therefore, we calculated an average rating volume for each doctor again (the correlations can be found in Multimedia Appendix 1).

Finally, to isolate the effects of expert and patient ratings on the click-based dependent variables, we included doctor-specific control variables. We controlled for the number of doctor referrals. On the basis of 160 million Medicare referrals from 2009 to 2012, this number indicated how many patients were referred by other doctors to a specific doctor. Doctor referrals only appear on a profile when they are greater than 0, so we included a dummy variable for whether doctors had any referrals or not (0=no, 1=yes). Furthermore, we controlled for each doctor's specialty; the number of practices in which she or he is employed; and whether the doctor has a profile image (0=no, 1=yes), photo gallery (0=no, 1=yes), or Web-based booking system incorporated in the profile (0=no, 1=yes), as well as whether a doctor has a premium profile (0=no, 1=yes) on the platform.

Model Specification

Most doctors in our sample received few profile or call clicks during the data collection period, and a few doctors had many visitors. Thus, both dependent count variables are overdispersed ($mean_{profile\ clicks}$ 4.12, variance 85.89; $mean_{call\ clicks}$ 0.34, variance 3.03), and we therefore modeled both dependent variables with a negative binomial regression [68]. To account for systematic differences between different kinds of doctors, we included a random intercept for each of the 57 doctor specialties:

$$\log(\text{profile clicks}_i) = \alpha_0 + \alpha_j + \beta_1(\text{expert rating}_i) + \beta_2(\text{consumer rating}_i) + \beta_3(\text{consumer rating}_i \times \text{rating volume}_i) + \Omega X_i + \varepsilon_i,$$

where i indexes the doctor and j the doctor specialty, X_i is the vector of doctor-specific control variables, and ε_i is the error. The model specification for our second dependent measure, call clicks, is analogous. For our call clicks model, we controlled for all variables; for the profile clicks model, we only controlled for doctor specialty, practice count, profile image, and premium profile, as the other control variables only become visible to consumers once they have clicked on a doctor's profile (and not on the search results page).

Although the analysis of real clickstream data is useful for getting first insights into the effect of simultaneously provided expert and patient ratings, these are only descriptive in nature. As we are interested in the causal process, we complement our first study with an experimental study in a controlled setting. In study 2, we systematically manipulated expert and patient ratings as well as rating volume. As such, study 2 aimed at finding further support for the causal influence of the number of underlying patient opinions (H1) by replicating the findings from study 1 in an experimental setting.

Study 2: Experimental Study

Procedure and Sample

We employed a 2×2 between-subjects design, in which we manipulated the valences of the expert and patient rating (positive vs negative) and the number of individual opinions the patient rating is based on (high volume vs low volume). We included only manipulations of conflicting ratings (ie, expert positive and patient negative and vice versa-we refer to this variable as "positive source type," as it can be "patient positive" or "expert positive"). Our objective was to compare a very high and a very low rating volume to be able to effectively contrast the respective effects. Drawing from prior research [64], observations of a Dutch health care rating platform (independer.nl), and especially from our study 1 dataset, we set the high rating volume to 142 (among top 1% of number of underlying opinions in our dataset). We set the low rating volume to 3 so that the aggregated patient rating resembled more than just a single but still only very few opinions. We recruited 125 undergraduate students from a Dutch university in exchange for course credit and randomly assigned them to 1 of the 4 experimental conditions in a computer lab. We excluded 13 participants who did not recall that there was a patient and

XSL•FO
RenderX

expert star rating on the site. In our sample (n=112), participants had a mean age of 23.1 years, and 60.7% (68/112) were female.

Participants imagined that they were looking for a hospital to perform a surgery, so they consulted a rating platform to help make their decision (the manipulation can be found in Multimedia Appendix 2). On the following page, the evaluation of a fictional hospital, on a fictional health care rating platform, featured both an expert and a patient star rating. Those ratings were conflicting and either negative (1.5 out of 5 stars) or positive (4.5 out of 5 stars). The patient rating was either based on 3 or 142 individual opinions. Next, participants indicated their attitudes toward the hospital [69,70] and their usage intentions [71]. These variables correlated (r=.83, P<.01), so we combined them into a single evaluative index ("hospital evaluation"; alpha=.96). We also measured participants' evaluations of experts and patients with regard to their expertise (experts: alpha=.94; consumers: alpha=.83 [72]) and trustworthiness (experts: alpha=.95; consumers: alpha=.88 [73]). All these measures featured 5-point scales (the correlations can be found in Multimedia Appendix 1; the reliability of scales are presented in Multimedia Appendix 3).

Results

Study 1: Field Study

Table 1 shows the results of the regression models. The analyses for the profile clicks and call clicks models exclude 5 doctors (profile clicks model) and 1213 doctors (call clicks model) because of missing values for the predictors, resulting in a sample size of 5294 for the profile clicks model and 4086 for the call clicks model. In our 2 initial models (models 1 and 3), we found a significant positive effect of the expert rating on both profile clicks (beta=.13, P<.001) and call clicks (beta=.32, P<.001) and of the patient rating on profile clicks (beta=.03, P=.02) but not on call clicks (beta=.04, P=.76). The interactions of patient rating and rating volume (mean centered [74]) on profile clicks (beta=.01, P<.001) and call clicks (beta=.01, P<.001) were, however, significant. To investigate these interaction effects further and assess the effect of the patient rating on the dependent variables with different numbers of underlying individual ratings, we repeated the analyses with a mean-centered value for the rating volume plus 3 SDs (models 2 and 4). As expected, the effect of the patient rating increased with the number of individual reviews aggregated to create the rating. As we noted previously, the effects of patient ratings were only significant for profile but not for call clicks for an average rating volume (Yelp rating based on 9 opinions), but their effects were significant and comparable with those of expert ratings when the underlying number of patient opinions was 1 SD higher (Yelp rating based on 27 opinions; profile clicks beta=.14, P<.001; call clicks beta=.28, P=.03). The effects of expert and patient ratings were not significantly different from each other: profile clicks beta=.01, SE 0.04, P=.37; call clicks beta=−.05, SE 0.15, P=.38. When the patient rating was based on the average rating volume plus 3 SDs (Yelp rating based on 63 opinions), the effects on both dependent variables (profile clicks beta=.35, P<.001; call clicks beta=.75, P=.02) were even stronger. With the number of underlying patient opinions increasing to 63, the effect of the patient rating becomes (marginally) significantly stronger than the effect of the expert rating (difference profile clicks beta=.23, SE 0.07, P=.004; call clicks β=.43, SE 0.32, P=.10).

The results of the analysis of field data thus confirm rating volume (ie, the number of underlying patient reviews) to be an important moderator of the impact of patient versus expert advice. Specifically, we find that when confronted with both sources simultaneously, prospective patients tend to base their evaluations of a physician on expert rather than patient advice in case the patient advice is based on small numbers. However, when the group of patients substantially grows in size, prospective patients start to rely more on patients rather than the expert in making a decision, which supports H1. Thus, patients are indeed influenced more strongly by advice from other patients than by advice from experts, but this occurs only when the patient advice is based on a large number of individual opinions. This finding is in line with prior research stating that volume increases the perceived diagnosticity and the effect of Web-based reviews (eg, [28,29]).

XSL•FO
RenderX

Table 1. Regression results of study 1.

Independent variables	Dependent variable: log profile clicks				Dependent variable: log call clicks			
	Model 1		Model 2		Model 3		Model 4	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Expert rating	0.13[a]	0.02	0.13[a]	0.02	0.32[a]	0.06	0.32[a]	0.06
Patient rating × rating volume (centered at mean)	0.01[a]	0.00	—[b]	—	0.01[a]	0.01	—	—
Patient rating (at mean level of rating volume)	0.03[c]	0.01	—	—	0.04[a]	0.05	—	—
Rating volume	0.00[a]	0.00	—	—	−0.00[a]	0.00	—	—
Patient rating × rating volume (centered at mean + 3 SD)	—	—	0.01[a]	0.00	—	—	0.01[a]	0.01
Patient rating (at mean level of rating volume + 3 SD)	—	—	0.35[a]	0.07	—	—	0.75[a]	0.31
Rating volume	—	—	0.00[a]	0.00	—	—	−0.01[a]	0.00
Controls								
Doctor referrals	—	—	N/A	—	−0.00[a]	0.11	−0.00[a]	0.11
Doctor referrals count	—	—	N/A	—	0.00[b]	0.00	0.00[a]	0.00
Practice count	0.00[a]	0.01	0.00[b]	0.01	0.07[a]	0.02	0.07[a]	0.02
Profile image	0.03[a]	0.03	0.03[b]	0.03	−0.07[a]	0.10	−0.07[a]	0.10
Photo gallery	—	—	—	—	0.64[a]	0.15	0.64[a]	0.15
Web-based booking	0.19[a]	0.03	0.19[a]	0.03	0.05[a]	0.11	0.05[a]	0.11
Premium profile	2.03[a]	0.12	2.03[a]	0.12	0.98[a]	0.44	0.98[a]	0.44
Intercept	0.72[a]	0.07	84[a]	0.09	−3.34[a]	0.29	−3.45[a]	0.30
Log-likelihood	−12910.8	—	—	—	−2257.7	—	—	—
Akaike information criterion	25843.5	—	—	—	4543.4	—	—	—

[a]Significant at the $P<.01$ level.

[b]Not applicable.

[c]Significant at the $P<.05$ level.

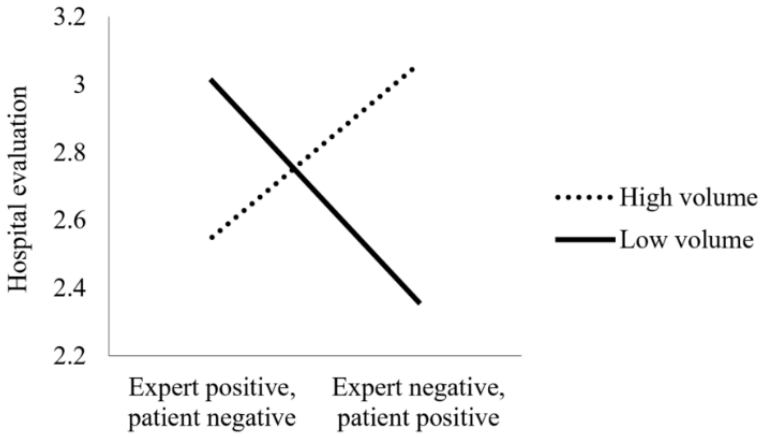
Study 2: Experimental Study

To check whether participants indeed perceived the experts and patients in our stimuli as such, we investigated whether expert and patient ratings were perceived differently with regard to expertise and trustworthiness, and whether this depended on the additional rating volume information. Experts were generally perceived as higher in expertise (mean 3.95, SD 0.73) than patients (mean 2.57, SD 0.69; $F_{1,110}=190.33$, $P<.001$), which was independent of the number of underlying patient opinions in the experimental condition ($P>.85$). Trustworthiness did not differ between experts and patients, again independent of the number of underlying patient opinions (all P values>.20).

The results of an analysis of variance indicated no significant main effects of positive source type (patient positive vs expert positive) and rating volume (142 vs 3) on hospital evaluation ($F_{\text{positive source type}1,108}=0.17$, $P=.68$; $F_{\text{volume}1,108}=0.52$, $P=.47$). However, as expected, we found a significant interaction effect of these 2 variables on hospital evaluation ($F_{1,108}=13.06$, $P<.001$). Further analyses demonstrated that when patient ratings are aggregated from a high number of opinions, prospective patients' evaluations are affected more strongly by patient than expert advice ($\text{mean}_{\text{patient positive/expert negative}}$ 3.06, SD 0.94; $\text{mean}_{\text{expert positive/patient negative}}$ 2.55, SD 0.89; $F_{1,108}=4.93$, $P=.03$). Conversely, when patient ratings are aggregated from a low volume, participants are affected more strongly by expert compared with patient advice ($\text{mean}_{\text{patient positive/expert negative}}$ 2.36, SD 0.76; $\text{mean}_{\text{expert positive/patient negative}}$ 3.01, SD 0.81; $F_{1,108}=8.42$, $P=.004$; Figure 1). Thus, study 2 finds further support for H1.

XSL•FO
RenderX

Figure 1. Study 2: Interaction effect of positive source type and rating volume on hospital evaluation.



The results of study 2 thus further support the findings from study 1: prospective patients are influenced more strongly by other patients when the patient evaluation is based on a larger number of individual opinions, but not when it is based on only a few observations. When the patient rating has been aggregated from only a small number of individual opinions, website users are instead more inclined to follow the expert advice. However, we find that participants still perceive experts to be higher in expertise, even than a large number of other patients. Moreover, and in contrast to prior findings [75], website users do not seem to trust experts less than other patients or assume any biases or ulterior motives. Consequently, prospective patients seem to be aware of the fact that a large number of other patients still lack the expertise to judge the objective outcome quality of a health service and also do not trust them more than an expert giving a rating. Yet, they are more inclined to follow their peers' compared with an expert's advice when choosing a health service.

Discussion

Principal Findings

This study examined how the evaluations and usage intentions of a health service is influenced by the simultaneous exposure to conflicting patient versus expert ratings and how this relationship is moderated by rating volume. The findings support the hypothesis that with increasing numbers of underlying patient opinions, the influence of the patient over the expert rating increases. In case the patient advice is based on small numbers, prospective patients tend to base their evaluations on an expert rather than patient advice. However, when the group of patients substantially grows in size, prospective patients start to rely more on patients rather than experts in making a decision. These patients are considered to be more similar yet less knowledgeable than experts. Thus, in line with social influence and source effects literature, patients are influenced more strongly by advice from other patients than by advice from experts but only when the patient advice is based on a large number of individual opinions. This has important implications for practitioners who design health care rating platforms. These implications hold special importance for how and which ratings to present to prospective patients to aid their decision-making process. It further helps us to understand the psychological processes underlying patients' preferences for advice from other patients or experts.

To the best of our knowledge, this study is the first to analyze the impact of the *simultaneous* presence of conflicting ratings from patient and expert sources on patient decision making. Research that provides participants with 1 source offers conflicting evidence for the preference for peer over expert advice (eg, [28,50]), but such studies cannot explicate the disaggregated effects of advice from different sources on a receiver or how consumers decide "which electronic word-of-mouth messages to adopt and which ones to reject" [76]. In 1 field study and 1 experimental study, we shed light on the differential effects of patient and expert advice in a health care context.

Limitations

In providing initial insights into the use of conflicting advice from different sources in patient decision-making, this study features several limitations that provide directions for further research. First, in our experimental study, we did not specify who the "experts" were. Prior research indicates that users mainly look for authority cues when assessing a website's credibility, but they rarely investigate who the expert sources actually are [44]. The participants in our studies perceived the experts as such. However, it would be interesting to explore whether the findings change when the advice comes from various experts, such as family doctors or specialist physicians.

Second, our findings only lay the foundation for understanding how users make a decision when confronted with different sources of information simultaneously. Although our analysis of clickstream data is high in external validity, it only produces correlational results. The sample was further not taken at random and should thus be interpreted with caution. Study 2 can likewise not quantify the exact effects of expert and patient ratings. Future research should investigate this further by making users choose between different offerings. Doing so will enable us to quantify the relative importance of different pieces of information.

Third, we did not analyze the potential moderating roles of other rating and platform characteristics. Platforms tend to feature both numerical ratings and textual reviews [77]. Other important characteristics, including language use, salience of the valence,

XSL•FO
RenderX

or further information about the sources and their reputation [76] thus deserve further inquiry.

Fourth, this study investigated Web-based ratings, one of many potential information sources a consumer employs to make health service decisions. Offline sources such as friends and family also strongly influence health service choice (eg, [78]). Further research might examine the differential effects of offline versus Web-based word of mouth and their interaction across the different stages of the decision-making process.

Conclusions

First, this study enhances the understanding of patients' use of Web-based decision support tools. More specifically, we shed light on the integration of simultaneously provided information from other patients and experts. We found that the opinion of a moderate number of other patients strongly influenced users' evaluations and choices of physicians, even overruling the conflicting opinions of experts. In this sense, choosing a health care provider does not seem to differ much from purchase choices for (for example) movies or restaurants. Even for credence services such as health care, others' subjective ratings of service experiences have a greater impact on decision making than more objective ratings of the service. Although the subjective patient experience is also important, it often fails to acknowledge the actual outcome quality of the service which patients have difficulties evaluating even after it has been performed [79]. These 2 aspects should be considered complementary for well-informed decisions, but our study suggests that this combination is not the route most consumers take when considering different sources of quality information on a Web-based rating platform. Even in health care contexts, users turn to consumers over expert advice, and this negligence of objective outcome quality information might lead them to make suboptimal choices, with considerable consequences for their health and society at large.

Second, for rating platform providers, our results support the trend of providing advice from both expert and patient sources. Experts and patients differ in their expertise and access to information, so they often emphasize different features in their judgments. These 2 sources of advice therefore should be regarded as complementary instead of substitutive input. Furthermore, the platform providers need to include the number of individual opinions underlying a patient rating because users might assume a rating is based on many opinions if the information is not evident. When this information is present, it acts as an important moderator, preventing an overemphasis on a collection of just a few patient ratings.

Acknowledgments

This work was supported by the Marketing Science Institute (Grant #4-1863). The authors thank Lois Sicking for her constructive input and support for this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The correlations of study 1 and 2.

[PDF File (Adobe PDF File), 152KB - jmir_v21i6e12454_app1.pdf]

Multimedia Appendix 2

Screenshots of a fictional Web-based health care rating platform used as the manipulations in study 2.

[PDF File (Adobe PDF File), 182KB - jmir_v21i6e12454_app2.pdf]

Multimedia Appendix 3

The reliability of scales.

[PDF File (Adobe PDF File), 132KB - jmir_v21i6e12454_app3.pdf]

References

1. Pan Y, Zhang JQ. Born unequal: a study of the helpfulness of user-generated product reviews. J. Retail 2011 Dec;87(4):598-612 [FREE Full text] [doi: 10.1016/j.jretai.2011.05.002]
2. Sillence E, Briggs P, Harris PR, Fishwick L. How do patients evaluate and make use of online health information? Soc Sci Med 2007 May;64(9):1853-1862 [FREE Full text] [doi: 10.1016/j.socscimed.2007.01.012] [Medline: 17328998]
3. DeAndrea DC, van der Heide B, Vendemia MA, Vang MH. How people evaluate online reviews. J Commun Res 2015;45(5):719-736 [FREE Full text] [doi: 10.1177/0093650215573862]
4. Grasso KL, Bell RA. Understanding health information seeking: a test of the risk perception attitude framework. J Health Commun 2015;20(12):1406-1414 [FREE Full text] [doi: 10.1080/10810730.2015.1018634] [Medline: 26161622]

XSL•FO
RenderX

5. Verlegh PW, Ryu G, Tuk MA, Feick L. Receiver responses to rewarded referrals: the motive inferences framework. J Acad Mark Sci 2013 Jan 29;41(6):669-682 [FREE Full text] [doi: 10.1007/s11747-013-0327-8]
6. Walther JB, Jang JW, Hanna Edwards AA. Evaluating health advice in a web 2.0 environment: the impact of multiple user-generated factors on HIV advice perceptions. Health Commun 2018;33(1):57-67 [FREE Full text] [doi: 10.1080/10410236.2016.1242036] [Medline: 27911096]
7. 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]
8. Liang B, Scammon DL. Incidence of online health information search: a useful proxy for public health risk perception. J Med Internet Res 2013 Jun 17;15(6):e114 [FREE Full text] [doi: 10.2196/jmir.2401] [Medline: 23773974]
9. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. J Am Med Assoc 2014 Feb 19;311(7):734-735 [FREE Full text] [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
10. Li S, Feng B, Chen M, Bell RA. Physician review websites: effects of the proportion and position of negative reviews on readers' willingness to choose the doctor. J Health Commun 2015 Apr;20(4):453-461 [FREE Full text] [doi: 10.1080/10810730.2014.977467] [Medline: 25749406]
11. McLennan S, Strech D, Reimann S. Developments in the frequency of ratings and evaluation tendencies: a review of German physician rating websites. J Med Internet Res 2017;19(8):e299 [FREE Full text] [doi: 10.2196/jmir.6599] [Medline: 28842391]
12. Emmert M, Halling F, Meier F. Evaluations of dentists on a German physician rating website: an analysis of the ratings. J Med Internet Res 2015 Jan 12;17(1):e15 [FREE Full text] [doi: 10.2196/jmir.3830] [Medline: 25582914]
13. Okike K, Peter-Bibb TK, Xie KC, Okike ON. Association between physician online rating and quality of care. J Med Internet Res 2016 Dec 13;18(12):e324 [FREE Full text] [doi: 10.2196/jmir.6612] [Medline: 27965191]
14. Rau J. Kaiser Health News. 2013. Nearly 1,500 Hospitals Penalized Under Medicare Program Rating Quality URL: https://khn.org/news/value-based-purchasing-medicare/ [WebCite Cache ID 6yS5LxbFr]
15. Medical Group Management Association. 2014. MGMA Survey: Practices Report Gradual Shifts Toward Provider Pay Tied to Quality and Patient Satisfaction Metrics URL: https://www.healthitoutcomes.com/doc/mgma-survey-shifts-quality-patient-satisfaction-metrics-0001 [WebCite Cache ID 6yRpRfx4r]
16. Berry LL, Bendapudi N. Health care: a fertile field for service research. J Serv Res 2007;10(2):111-122 [FREE Full text] [doi: 10.1177/1094670507306682]
17. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011 Dec 6;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]
18. Adams JG, Biros MH. The elusive nature of quality. Acad Emerg Med 2002 Nov;9(11):1067-1070 [FREE Full text] [doi: 10.1197/aemj.9.11.1067] [Medline: 12414452]
19. Healthgrades. 2014. Doctors, hospitals and health plans: how consumers choose URL: http://www.multivu.com/assets/63942/documents/63942-healthgrades-infographic-original.pdf [WebCite Cache ID 6yRrR9ZBb]
20. Büyüközkan G, Çifçi G, Güleryüz S. Strategic analysis of healthcare service quality using fuzzy AHP methodology. Expert Syst Appl 2011 Aug;38(8):9407-9424 [FREE Full text] [doi: 10.1197/aemj.9.11.1067]
21. McColl-Kennedy JR, Vargo SL, Dagger TS, Sweeney JC, van Kasteren Y. Health care customer value cocreation practice styles. J Serv Res 2012 May;15(4):370-389 [FREE Full text] [doi: 10.1177/1094670512442806]
22. Gray B, Vandergrift J, Lipner RS, Gao G, McCullough J. Evaluating Website Ratings: How Well do Patient Ratings of Physicians Predict Their quality of Care? In: Annual Research Meeting. 2014 Presented at: AcademyHealth; June 2014; Washington, DC.
23. 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 Nov 16;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
24. Aiken KD, Boush DM. Trustmarks, objective-source ratings, and implied investments in advertising: investigating online trust and the context-specific nature of internet signals. J Acad Mark Sci 2006 Jul 1;34(3):308-323 [FREE Full text] [doi: 10.1177/0092070304271004]
25. Chevalier J, Mayzlin D. The effect of word of mouth on sales: online book reviews. J Market Res 2018 Oct 10;43(3):345-354 [FREE Full text] [doi: 10.1509/jmkr.43.3.345]
26. Lagu T, Hannon N, Rothberg M, Lindenauer P. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
27. Zimmermann M, Jucks R. How experts' use of medical technical jargon in different types of online health forums affects perceived information credibility: randomized experiment with laypersons. J Med Internet Res 2018 Dec 23;20(1):e30 [FREE Full text] [doi: 10.2196/jmir.8346] [Medline: 29362212]
28. Flanagin AJ, Metzger MJ. Trusting expert- versus user-generated ratings online: the role of information volume, valence, and consumer characteristics. Comput Human Behav 2013;29(4):1626-1634 [FREE Full text] [doi: 10.1016/j.chb.2013.02.001]

XSL•FO
RenderX

29. Khare A, Labrecque LI, Asare AK. The assimilative and contrastive effects of word-of-mouth volume: an experimental examination of online consumer ratings. J. Retail 2011 Mar;87(1):111-126 [FREE Full text] [doi: 10.1016/j.jretai.2011.01.005]
30. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. J Med Internet Res 2013 Feb 01;15(2):e24 [FREE Full text] [doi: 10.2196/jmir.2360] [Medline: 23372115]
31. Gershoff AD, Broniarczyk SM, West PM. Recommendation or evaluation? task sensitivity in information source selection. J Consum Res 2001 Dec;28(3):418-438 [FREE Full text] [doi: 10.1086/323730]
32. Gilly MC, Graham JL, Wolfinbarger MF, Yale LJ. A dyadic study of interpersonal information search. J Acad Mark Sci 1998;26(2):83-100 [FREE Full text] [doi: 10.1177/0092070398262001]
33. Ketelaar PE, Willemsen LM, Sleven L, Kerkhof P. The good, the bad, and the expert: how consumer expertise affects review valence effects on purchase intentions in online product reviews. J Comput Mediat Commun 2015 Nov 27;20(6):649-666 [FREE Full text] [doi: 10.1111/jcc4.12139]
34. Brock TC. Communicator-recipient similarity and decision change. J Pers Soc Psychol 1965;1(6):650-654 [FREE Full text] [doi: 10.1037/h0022081]
35. Simons HW, Berkowitz NN, Moyer RJ. Similarity, credibility, and attitude change: a review and a theory. Psychol Bull 1970;73(1):1-16 [FREE Full text] [doi: 10.1037/h0028429]
36. Ruef M, Aldrich HE, Carter NM. The structure of founding teams: homophily, strong ties, and isolation among U.S. entrepreneurs. Am Sociol Rev 2003 Apr;68(2):195-222 [FREE Full text] [doi: 10.2307/1519766]
37. Price LL, Feick LF. The role of interpersonal sources in external search: an informational perspective. Adv Consum Res 1984;11:250-255 [FREE Full text]
38. Wang Z, Walther JB, Pingree S, Hawkins RP. Health information, credibility, homophily, and influence via the internet: web sites versus discussion groups. Health Commun 2008 Jul;23(4):358-368 [FREE Full text] [doi: 10.1080/10410230802229738] [Medline: 18702000]
39. Goldstein NJ, Cialdini RB, Griskevicius V. A room with a viewpoint: using social norms to motivate environmental conservation in hotels. J Consum Res 2008;35(3):472-482 [FREE Full text] [doi: 10.1086/586910]
40. Alba JW, Hutchinson JW. Dimensions of consumer expertise. J Consum Res 1987 Mar;13(4):411-454 [FREE Full text] [doi: 10.1086/209080]
41. Gotlieb JB, Sarel D. Comparative advertising effectiveness: the role of involvement and source credibility. J Advert 1991;20(1):38-45 [FREE Full text] [doi: 10.1080/00913367.1991.10673205]
42. Brown J, Broderick AJ, Lee N. Word of mouth communication within online communities: conceptualizing the online social network. J Interact Mark 2007;21(3):2-20 [FREE Full text] [doi: 10.1002/dir.20082]
43. Eastin MS. Credibility assessments of online health information: the effects of source expertise and knowledge of content. J Comput Mediat Commun 2001;6(4):- [FREE Full text] [doi: 10.1111/j.1083-6101.2001.tb00126.x]
44. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. Br Med J 2002 Mar 09;324(7337):573-577 [FREE Full text] [doi: 10.1136/bmj.324.7337.573] [Medline: 11884321]
45. Hu Y, Shyam Sundar S. Effects of online health sources on credibility and behavioral intentions. Communic Res 2010 Nov 25;37(1):105-132 [FREE Full text] [doi: 10.1177/0093650209351512]
46. Mitchell AA, Dacin PA. The assessment of alternative measures of consumer expertise. J Consum Res 1996;23(3):219-239. [doi: 10.1086/209479]
47. Bansal HS, Voyer PA. Word-of-mouth processes within a services purchase decision context. J Serv Res 2000;3(2):166-177 [FREE Full text] [doi: 10.1177/109467050032005]
48. Schweiger S, Oeberst A, Cress U. Confirmation bias in web-based search: a randomized online study on the effects of expert information and social tags on information search and evaluation. J Med Internet Res 2014 Mar 26;16(3):e94 [FREE Full text] [doi: 10.2196/jmir.3044] [Medline: 24670677]
49. Feng B, MacGeorge EL. The influences of message and source factors on advice outcomes. Commun Res 2010 Jun 09;37(4):553-575 [FREE Full text] [doi: 10.1177/0093650210368258]
50. Zhang Z, Ye Q, Law R, Li Y. The impact of e-word-of-mouth on the online popularity of restaurants: a comparison of consumer reviews and editor reviews. Int J Hosp Manag 2010;29(4):694-700 [FREE Full text] [doi: 10.1016/j.ijhm.2010.02.002]
51. Feldman SP, Spencer MC. The effect of personal influence in the selection of consumer services. In: Holloway RJ, Mittelstaedt RA, Venkatesan M, editors. Consumer Behavior: Contemporary Research in Action. Boston: Houghton Mifflin; 1971:247-257.
52. Woodside AG, Davenport Jr JW. The effect of salesman similarity and expertise on consumer purchasing behavior. J Mark Res 1974 May;11(2):198-202 [FREE Full text] [doi: 10.2307/3150562]
53. Floyd K, Freling R, Alhoqail S, Cho HY, Freling T. How online product reviews affect retail sales: a meta-analysis. J Retail 2014;90(2):217-232 [FREE Full text] [doi: 10.1016/j.jretai.2014.04.004]
54. Lim YS, van der Heide B. Evaluating the wisdom of strangers: the perceived credibility of online consumer reviews on Yelp. J Comput Mediat Comm 2015;20(1):67-82 [FREE Full text] [doi: 10.1111/jcc4.12093]

XSL•FO
RenderX

55. van der Heide B, Lim YS. On the conditional cueing of credibility heuristics. Communic Res 2016;43(5):672-693 [FREE Full text] [doi: 10.1177/0093650214565915]
56. Cialdini RB, Goldstein NJ. Social influence: compliance and conformity. Annu Rev Psychol 2004;55:591-621 [FREE Full text] [doi: 10.1146/annurev.psych.55.090902.142015] [Medline: 14744228]
57. Tanford S, Penrod S. Social influence model: a formal integration of research on majority and minority influence processes. Psychol Bull 1984;95(2):189-225 [FREE Full text] [doi: 10.1037/0033-2909.95.2.189]
58. López-Pintado D, Watts DJ. Social influence, binary decisions and collective dynamics. Ration Soc 2008 Nov;20(4):399-443 [FREE Full text] [doi: 10.1177/1043463108096787]
59. Watts DJ, Dodds PS. Influentials, networks, and public opinion formation. J Consum Res 2007 Dec;34(4):441-458 [FREE Full text] [doi: 10.1086/518527]
60. Hasher L, Zacks RT. Automatic processing of fundamental information: the case of frequency of occurrence. Am Psychol 1984;39(12):1372-1388 [FREE Full text] [doi: 10.1037/0003-066X.39.12.1372]
61. Baker SM, Petty RE. Majority and minority influence: source-position imbalance as a determinant of message scrutiny. J Pers Soc Psychol 1994;67(1):5-19 [FREE Full text] [doi: 10.1037/0022-3514.67.1.5]
62. Latané B. The psychology of social impact. Am Psychol 1981;36(4):343-356 [FREE Full text] [doi: 10.1037/0003-066X.36.4.343]
63. Dholakia UM, Basuroy S, Soltysinski K. Auction or agent (or both)? A study of moderators of the herding bias in digital auctions. Int J Res Market 2002 Jun;19(2):115-130 [FREE Full text] [doi: 10.1016/S0167-8116(02)00064-2]
64. Liu Y. Word of mouth for movies: its dynamics and impact on box office revenue. J Mark 2006 Jul;70(3):74-89 [FREE Full text] [doi: 10.1509/jmkg.70.3.74]
65. Salganik MJ, Dodds PS, Watts DJ. Experimental study of inequality and unpredictability in an artificial cultural market. Science 2006 Feb 10;311(5762):854-856 [FREE Full text] [doi: 10.1126/science.1121066] [Medline: 16469928]
66. Prendergast C, Stole L. Impetuous youngsters and jaded old-timers: acquiring a reputation for learning. J Polit Econ 1996 Dec;104(6):1105-1134 [FREE Full text] [doi: 10.1086/262055]
67. Ye Q, Law R, Gu B, Chen W. The influence of user-generated content on traveler behavior: an empirical investigation on the effects of e-word-of-mouth to hotel online bookings. Comput Hum Behav 2011;27(2):634-639 [FREE Full text] [doi: 10.1016/j.chb.2010.04.014]
68. Greene WH. Econometric Analysis 5th Edition. Upper Saddle River, NJ: Prentice Hall; 2003.
69. Becker-Olsen KL. And now, a word from our sponsor--a look at the effects of sponsored content and banner advertising. J Advert 2003;32(2):17-32 [FREE Full text] [doi: 10.1080/00913367.2003.10639130]
70. Rodgers S. The effects of sponsor relevance on consumer reactions to internet sponsorships. J Advert 2003;32(4):67-76. [doi: 10.1080/00913367.2003.10639141]
71. Chandran S, Morwitz VG. Effects of participative pricing on consumers' cognitions and actions: a goal theoretic perspective. J Consumer Res 2005 Sep;32(2):249-259 [FREE Full text] [doi: 10.1086/432234]
72. Ohanian R. Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. J Advert 1990 Oct;19(3):39-52 [FREE Full text] [doi: 10.1080/00913367.1990.10673191]
73. Newell SJ, Goldsmith RE. The development of a scale to measure perceived corporate credibility. J Bus Res 2001 Jun;52(3):235-247 [FREE Full text] [doi: 10.1016/S0148-2963(99)00104-6]
74. Jaccard J, Turrisi R. Interaction Effects in Multiple Regression. 2nd Edition. Thousand Oaks, CA: Sage Publications; 2003.
75. Willemsen L, Neijens P, Bronner F. The ironic effect of source identification on the perceived credibility of online product reviewers. J Comput-Mediat Comm 2012 Oct 10;18(1):16-31 [FREE Full text] [doi: 10.1111/j.1083-6101.2012.01598.x]
76. King RA, Racherla P, Bush VD. What we know and don't know about online word-of-mouth: a review and synthesis of the literature. J Interact Mark 2014;28(3):167-183 [FREE Full text] [doi: 10.1016/j.intmar.2014.02.001]
77. Emmert M, Meszmer N, Sander U. Do health care providers use online patient ratings to improve the quality of care? Results from an online-based cross-sectional study. J Med Internet Res 2016;18(9):e254 [FREE Full text] [doi: 10.2196/jmir.5889] [Medline: 27644135]
78. Hoerger TJ, Howard LZ. Search behavior and choice of physician in the market for prenatal care. Med Care 1995;33(4):332-349. [Medline: 7731276]
79. Darby MR, Karni E. Free competition and the optimal amount of fraud. J Law Econ 1973;16(1):67-88 [FREE Full text] [doi: 10.1086/466756]

XSL•FO
RenderX

Edited by G Eysenbach; submitted 09.10.18; peer-reviewed by E Van Den Broek-Altenburg, E Micaiah; comments to author 27.12.18; revised version received 13.04.19; accepted 20.05.19; published 26.06.19.

Please cite as:
Kranzbühler AM, Kleijnen MHP, Verlegh PWJ, Teerling M
When Similarity Beats Expertise—Differential Effects of Patient and Expert Ratings on Physician Choice: Field and Experimental Study
J Med Internet Res 2019;21(6):e12454
URL: http://www.jmir.org/2019/6/e12454/
doi:10.2196/12454
PMID:31244481

©Anne-Madeleine Kranzbühler, Mirella H P Kleijnen, Peeter W J Verlegh, Marije Teerling. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Impact of Web-Based Ratings on Patient Choice of a Primary Care Physician Versus a Specialist: Randomized Controlled Experiment

Siyue Li[1], PhD; Austin Hubner[2], MA

[1]College of Media and International Culture, Zhejiang University, Hangzhou, China

[2]School of Communication, The Ohio State University, Columbus, OH, United States

Corresponding Author:
Siyue Li, PhD
College of Media and International Culture
Zhejiang University
Tianmushan Rd #148
Hangzhou, 310000
China
Phone: 86 87951596
Email: siyueaprilli@gmail.com

Abstract

Background: Physician review websites have empowered prospective patients to acquire information about physicians. However, little is known about how Web-based ratings on different aspects of a physician may affect patients' selection of physicians differently.

Objective: The objectives of this study were to examine (1) how patients weigh ratings on a physician's technical skills and interpersonal skills in their selection of physicians and (2) whether and how people's choice of a primary care physician versus a specialist is affected differently by Web-based ratings.

Methods: A 2×2×2×2 between-subjects experiment was conducted. Over 600 participants were recruited through a crowdsourcing website and randomly assigned to view a mockup physician review Web page that contained information on a physician's basic information and patients' ratings. After reviewing the Web page, participants were asked to complete a survey on their perceptions of the physician and willingness to seek health care from the physician.

Results: The results showed that participants were more willing to choose a physician with higher ratings on technical skills than on interpersonal skills compared with a physician with higher ratings on interpersonal skills than on technical skills, $t_{369.96}$=22.36, P<.001, Cohen d=1.22. In the selection of different types of physicians, patients were more likely to choose a specialist with higher ratings on technical skills than on interpersonal skills, compared with a primary care physician with the same ratings, $F_{1,521}$=5.34, P=.021.

Conclusions: The findings suggest that people place more weight on technical skills than interpersonal skills in their selection of a physician based on their ratings on the Web. Specifically, people are more likely to make a compromise on interpersonal skills in their choice of a specialist compared with a primary care physician. This study emphasizes the importance of examining Web-based physician ratings in a more nuanced way in relation to the selection of different types of physicians.

Trial Registration: ISRCTN Registry ISRCTN91316463; http://www.isrctn.com/ISRCTN91316463

(J Med Internet Res 2019;21(6):e11188) doi:10.2196/11188

KEYWORDS

technical skills; interpersonal skills; physician ratings; physician selection

XSL•FO
RenderX

Introduction

Background

The role of patients in health care is undergoing a remarkable transition. Although traditional patients took a passive role in their health care, modern patients are actively involved in health decision making [1,2]. For instance, patients are increasingly turning to Web-based physician review websites (PRWs) to learn information about their physicians [3,4]. Indeed, a national survey suggests that more than half of the respondents consider PRWs as an important source for information when choosing a physician [3]. These websites not only provide information about a physician's basic information and qualifications but also present peer-to-peer consumer reviews of the physician. On the basis of the reviews, health consumers are able to learn about other patients' experiences, concerns, and levels of satisfaction about a specific physician.

With the growing popularity of PRWs, researchers have recently begun to examine the role of these websites in people's health decision making [5-8]. The scholarship on PRWs covers a wide range of topics, including but not limited to demographics of website users, structures of the portals, patterns of website usage, and content of reviews [9-11]. Technical skills and interpersonal skills reside at the core of a physician's qualifications and are commonly rated on PRWs [12,13]. However, little is known about how ratings on these different aspects of a physician may affect patients' choice differently [14]. Previous research presents mixed results on how people set the priority of technical and interpersonal skills in physician selection [15,16]. Therefore, the first goal of this study was to examine how patients prioritize technical and interpersonal skills in their physician selection based on ratings on PRWs.

The second objective of this study was to examine whether and how people's choice of a primary care physician versus a specialist is affected differently by Web-based ratings. In the United States, patients are allowed greater autonomy to choose their primary care physicians compared with specialists. As a result, significantly more research focuses on people's selection of primary care physicians than physicians of other types [16-18]. Nowadays, however, patients are more involved in the choice of specialists in part because of the easy access of health information on the Web. Given that primary care physicians and specialists take on different roles in health care, patients may apply different criteria to select different types of physicians [3]. In this study, we specifically examine how ratings on a physician's technical and interpersonal skills may affect patients' choice of primary care physicians and specialists differently. We only examine medical doctors and exclude dentists because medicine and dentistry are often considered separately in health care. They involve different education systems, physician networks, medical records, and payment systems. Some PRWs list dentist as a different category from medical doctors and other sites might not include dentists.

Technical Versus Interpersonal Skills

In the era of health consumerism, people tend to evaluate a wide array of factors in their selection of physicians. Research on physician selection criteria has shown that people not only consider the factors pertaining to a physician (eg, sex, age, race, and qualification) but also evaluate many other contextual and economic factors, such as office location and insurance coverage [12,16,18].

Despite a large collection of factors to consider, technical skills and interpersonal skills are central to the evaluation of a physician's qualifications [12,13]. Technical skills concern medical knowledge and expertise in a physician's area. Overall, patients prefer a physician who is skilled and knowledgeable in their domain of expertise with the ability to provide accurate diagnosis and treatment [12]. Interpersonal skills center on the communication style of a physician [19]. Especially with the recent push for patient-centered care, physicians of all types are reevaluating their approach to communicating with patients [20,21]. Patients, in general, prefer a physician who is easy to talk to and willing to listen [22]. Physicians with a caring and friendly style lead to high levels of patient satisfaction [23].

Although both technical and interpersonal skills are important considerations in patients' choice of physicians, it is unclear how patients set the priority of the 2 factors. Research has presented inconsistent findings regarding patients' selection criteria of a physician [15,16,24,25]. A body of literature focusing on primary care physicians found that people expressed a clear preference for technical skills over interpersonal skills [12,26]. However, other studies show that communication skills are the most important determinant in patients' choice of a primary care physician [15,16,22]. Various aspects of interpersonal skills, such as a caring attitude and responsiveness, are found to be preferred over technical skills in people's choice of a primary care physician [16,27].

Similarly, research on patients' selection of specialists has generated mixed findings on patients' preferences on technical skills versus interpersonal skills [13,24,25]. Hoerger and Howard [25] found that women rated medical expertise as the leading reason for choice of a prenatal care physician. On the contrary, Dunlea and Lenert [24] surveyed people over their preference of a specialist with a hypothetical referral of an asymptomatic condition and concluded that communication skills and the shared decision-making style were considered more important than medical expertise in their choice.

PRWs provide prospective patients with valuable information regarding physicians. The valence of Web-based reviews tends to affect people's perceptions of physicians and their intention to choose the physicians [10]. The decision is even more complicated when people are exposed to reviews with opposite valence in different skills. For example, physicians may receive positive feedback on their technical competence but negative or neutral reviews on their interpersonal skills (or vice versa). Under such circumstances, patients may need to make tradeoffs between a physician's technical and interpersonal skills. As previous research presented conflicting findings on patients' preference over a physician's technical skills versus interpersonal skills, the following research question was raised:

> *RQ1: Are people more willing to choose a physician with higher ratings on technical skills than on interpersonal skills, or a physician with higher ratings on interpersonal skills than on technical skills?*

XSL•FO
RenderX

Physician Types and Skills

Extant literature has demonstrated the importance of technical skills and interpersonal skills in people's consideration of a physician, regardless of specialty [13]. However, the relative importance of technical and interpersonal skills may vary as a function of physician types. First, primary care physicians and specialists provide distinct medical services that require different levels of technical skills. In the United States, patients visit primary care physicians mainly for comprehensive care at the point of first contact, whereas they count on specialists for more specialized and advanced care. Primary care physicians generally have a wide range of medical knowledge and offer comprehensive care to patients, but they may not have advanced training for a particular domain of health. Specialists, as compared with primary care physicians, have advanced training in a particular branch of medicine (eg, bone or lung health) and are capable of providing more specialized care for patients. Patients who need advanced care for specific health conditions are usually referred by their primary care physicians to a specialist. Owing to the distinct services provided by primary care physicians and specialists, patients' expectations for the technical skills of their physicians may vary between the types. Technical skills are likely more important in the assessment of specialists compared with primary care physicians.

Beyond differing responsibilities in health care, primary care physicians and specialists differ in their relationships with patients. Long-term relationships are typically expected with primary care physicians whereas specialist-patient relationships are largely bound by specific health problems and in many cases are short-lived [13,28]. Many patients choose to stay with the same primary care physician for an extended period of time. These relationships, for instance, might last over 10 years [29]. On the contrary, patients who are having a specific health condition treated may not need to keep a long-term relationship with their specialists. As sophisticated interpersonal skills are found to be important factors for building long-term relationships [30], patients may place more weight on interpersonal skills in their selection of a primary care physician compared with a specialist.

A limited body of literature has compared the relative importance of technical and interpersonal skills in people's choice of primary care physicians and specialists. Hanna et al [31] found that communication skills were the leading factor in the selection of a primary care physician, whereas medical expertise was the key determinant in the choice of a specialist. In addition, a study on female patients' selection of a primary care physician versus a specialist showed that people rated medical expertise to be a more important determinant in selection of a surgeon compared with a primary care physician [13]. Although we know little about the comparison between primary care physicians and specialists based on their skills, the sparse research evidence indicates that technical skills are likely to be more valued in the selection of specialists and interpersonal skills to be more valued in the selection of primary care physicians. When physicians receive Web-based ratings that indicate different qualities of their technical and interpersonal skills, patients may need to make a tradeoff. It is expected that patients are more willing to sacrifice interpersonal skills for technical skills in their selection of a specialist compared with a primary care physician. In contrast, patients are more likely to make a compromise on technical skills than interpersonal skills in their selection of a primary care physician versus a specialist. The following hypotheses were proposed:

> *H1: People are more willing to choose a specialist who has higher ratings on technical skills than on interpersonal skills, compared with a primary care physician who has the same ratings.*
>
> *H2: People are more willing to choose a primary care physician who has higher ratings on interpersonal skills than on technical skills, compared with a specialist who has the same ratings.*

Methods

Ethical Approval

The Institutional Review Board at the Ohio State University approved all study procedures.

Sample

A total of 608 participants completed the Web-based experiment. Participants were recruited via the crowdsourcing website, Amazon's Mechanical Turk (mTurk), and compensated for their time. We excluded people who failed the attention checks (n=26) and those who spent no time (n=1) or less than 5 seconds on the Web page (n=41). Of the 540 valid cases, 300 (300/540, 55.6%) were male and 239 (239/540, 44.3%) were female, with an average age of 35.83 (SD 11.30) years. The majority of the participants indicated that they were Caucasian (402/40, 74.4%), followed by Asian/Asian American (60/540, 11.1%), Hispanic/Latino (32/540, 5.9%), and African American (31/540, 5.7%). Demographics are provided in Table 1.

XSL•FO
RenderX

Table 1.

Characteristics	Total	Experimental group							
		Primary care physician				Specialist			
		Moderate ratings on technical skills		High ratings on technical skills		Moderate ratings on technical skills		High ratings on technical skills	
		IS[a], Moderate (n=65)	IS, High (n=69)	IS, Moderate (n=70)	IS, High (n=69)	IS, Moderate (n=69)	IS, High (n=67)	IS, Moderate (n=67)	IS, High (n=64)
Gender, n (%)									
Male	300 (55.6)	39 (60)	33 (47.8)	39 (55.7)	39 (56.5)	37 (53.6)	32 (47.8)	43 (64.2)	38 (59.4)
Female	239 (44.3)	26 (40)	36 (52.2)	31 (44.3)	30 (43.5)	31 (44.9)	35 (52.2)	24 (35.8)	26 (40.6)
Unspecified	1 (0.2)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1.4)	0 (0)	0 (0)	0 (0)
Education, n (%)									
High school graduate or less	74 (13.7)	9 (13.8)	6 (8.7)	12 (17.1)	8 (11.6)	10 (14.5)	10 (14.9)	7 (10.4)	12 (18.8)
Some college	135 (25)	12 (18.5)	25 (36.2)	20 (28.6)	15 (21.7)	14 (20.3)	17 (25.4)	18 (26.9)	14 (21.9)
2-years degree	84 (15.6)	13 (20)	5 (7.2)	11 (15.7)	11 (15.9)	12 (17.4)	9 (13.4)	12 (17.9)	11 (17.2)
Bachelor's degree	190 (35.2)	23 (35.4)	27 (39.1)	18 (25.7)	23 (33.3)	26 (37.7)	27 (40.3)	22 (32.8)	24 (37.5)
Graduate degree	57 (10.6)	8 (12.3)	6 (8.6)	9 (12.8)	12 (17.4)	7 (10.1)	4 (6.0)	8 (11.9)	3 (4.7)
Age (years), mean (SD)	35.8 (11.3)	35 (10.6)	34.6 (12.5)	36.5 (10.1)	37.2 (13.1)	33.8 (9.2)	36.4 (10.8)	37.1 (12.3)	36.0 (11.2)
Income, n (%)									
Less than 20,000	100 (18.5)	11 (16.9)	12 (17.4)	16 (22.9)	14 (20.3)	16 (23.2)	13 (19.4)	8 (11.9)	10 (15.6)
20,000 to <40,000	163 (30.2)	17 (26.1)	22 (31.9)	20 (28.6)	23 (33.3)	24 (34.8)	16 (23.9)	21 (31.3)	20 (31.2)
40,000 to <60,000	114 (21.1)	17 (26.1)	14 (20.3)	10 (14.3)	16 (23.2)	15 (21.7)	13 (19.4)	11 (16.4)	18 (28.1)
60,000 to <80,000	80 (14.8)	6 (9.2)	10 (14.4)	14 (20.0)	6 (8.7)	9 (13)	13 (19.4)	13 (19.4)	9 (14.1)
80,000 to <100,000	39 (7.2)	5 (7.7)	3 (4.3)	8 (11.4)	6 (8.7)	1 (1.4)	6 (9.0)	6 (9.0)	4 (6.3)
100,000 and higher	44 (8.1)	9 (13.8)	8 (11.6)	2 (2.9)	4 (5.8)	4 (5.8)	6 (9.0)	8 (11.9)	3 (4.7)
Race, n (%)									
Caucasian	402 (74.4)	52 (80)	51 (73.9)	52 (74.3)	46 (66.7)	55 (79.7)	50 (74.6)	50 (74.6)	46 (71.9)
Hispanic	32 (5.9)	1 (1.5)	2 (2.9)	4 (5.7)	4 (5.8)	3 (4.3)	5 (7.5)	4 (6.0)	9 (14.1)
African American	31 (5.7)	6 (9.2)	6 (8.7)	4 (5.7)	4 (5.8)	4 (5.8)	2 (3.0)	2 (3.0)	3 (4.7)
Native American/ Alaskan Native	5 (0.9)	0 (0)	0 (0)	0 (0)	1 (1.4)	2 (2.9)	0 (0)	2 (3.0)	0 (0)
Asian	60 (11.1)	5 (7.7)	9 (13)	7 (10)	14 (20.3)	3 (4.3)	9 (13.4)	8 (11.9)	5 (7.8)
Middle Eastern	2 (0.4)	1 (1.5)	0 (0)	1 (1.4)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Pacific Islander	1 (0.2)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (1.5)	0 (0)
Other	6 (1.1)	0 (0)	1 (1.4)	2 (2.9)	0 (0)	1 (1.4)	1 (1.5)	0 (0)	1 (1.6)

[a]IS: interpersonal skills

Research Design

To investigate the proposed research question and hypotheses, a 2 (ratings on interpersonal skills: high versus moderate) × 2 (ratings on technical skills: high versus moderate) × 2 (physician specialty: primary care physician versus specialist) × 2 (order of ratings: interpersonal skills first versus technical skills first) between-subjects factorial design was employed. Participants were randomly assigned to one of the 16 experimental conditions and instructed to read through a cover story describing a medical condition in which they need to find a new physician. They were then asked to view a mockup physician review page and complete a questionnaire about their perceptions of the reviewed physician and their willingness to choose the physician.

Stimulus Materials

Following consent, participants were presented a cover story to read. On the basis of the type of physician that they were assigned to, participants were asked to imagine themselves in a situation looking for either a primary care physician or a surgeon. The vignette about a primary care physician depicted a situation that the participant recently moved to a new city and

XSL•FO
RenderX

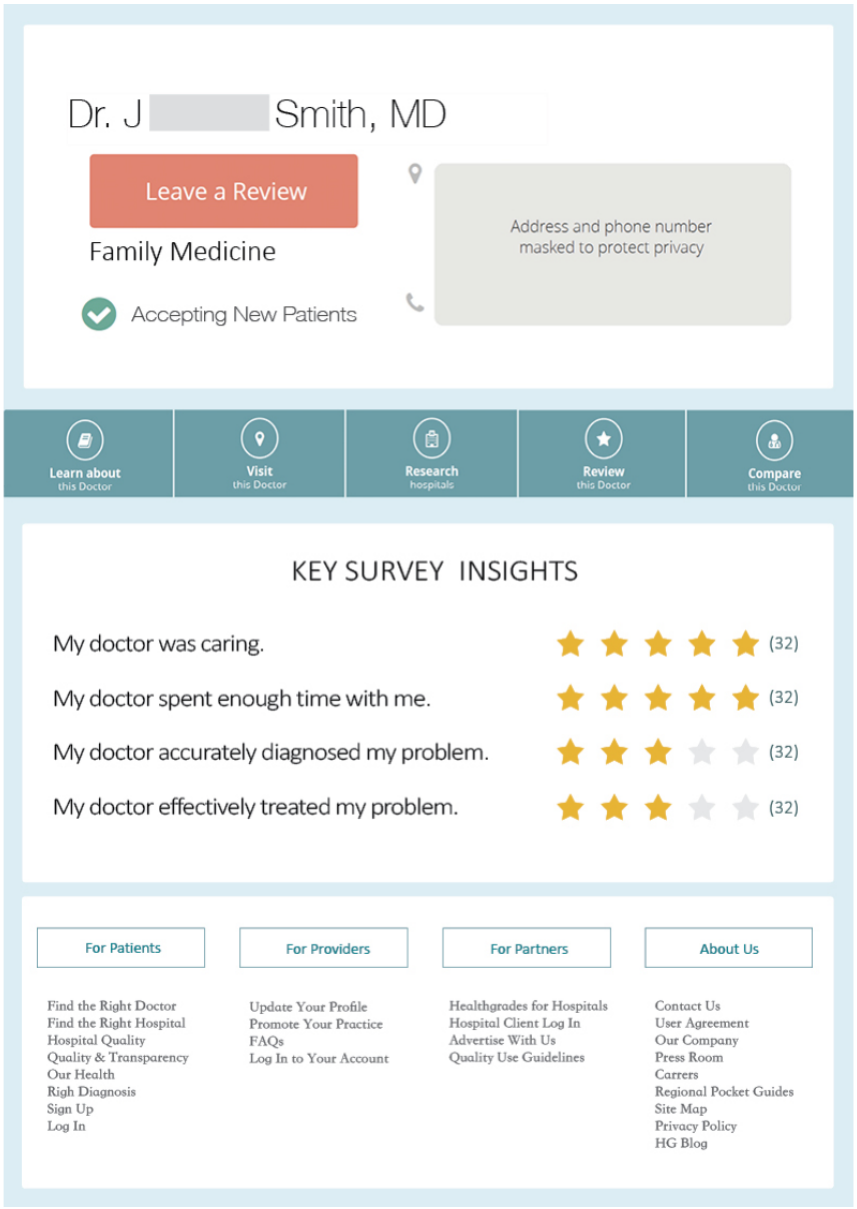
was in need of a new primary care physician. Owing to a lack of input from family members and friends, they decided to search for primary care physicians on PRWs. The vignette about a surgeon described a situation in which the participant had lasting back pains. The primary care physician suspected that the patient may need spinal surgery and provided a list of surgeons to choose from. The participant decided to search for the recommended surgeons on PRWs. After reading through the scenario and imaging themselves in the described situation, each participant was directed to a physician review page to learn about the physician.

A total of 16 physician review pages were developed for this study (see Figure 1). The top part of each page listed basic information about a physician, including the physician's name (Dr J Smith), the specialty (family medicine or surgeon), and information on new patient acceptance (accepting new patients). To manipulate the type of a physician, half of the Web pages listed the physician's specialty as family medicine and the other half described the physician as a surgeon.

Each page contained 4 aggregated rating categories about Dr Smith, including 2 items on technical skills ("My doctor accurately diagnosed my problem" and "My doctor effectively treated my problem") and 2 on interpersonal skills ("My doctor was caring" and "My doctor spent enough time with me"). Past research has suggested that a physician's skills on diagnosis and treatment are among the most important considerations when selecting a physician [32]. In addition, a physician's personal manner as well as time spent with a patient are critical to a patient's satisfaction on the physician's interpersonal skills [32,33]. These 4 categories frequently appear on PRWs [34] and thus are adopted in this study. To manipulate the valence of physician ratings, these rating categories were assigned different star ratings. Each rating category was presented in the form of aggregated ratings. In the conditions where a physician received high ratings on technical skills, the 2 items pertaining to technical skills were given 5/5-star ratings. In the conditions of moderate ratings on technical skills, the same items were assigned 3/5-star ratings. We chose to examine moderate ratings instead of low ratings in this study because research suggests that low ratings are relatively uncommon on PRWs . The valence of a physician's interpersonal skills was manipulated in the same way. Furthermore, the rating categories were presented to participants in counterbalanced order to control for the impact of rating order effects. In half of the experimental conditions, the 2 rating categories on technical skills were displayed before the 2 categories on interpersonal skills. In the other half, ratings on technical skills were presented beneath the ratings on interpersonal skills.

XSL•FO
RenderX

Figure 1. An example of the physician review page stimuli.



Measures

Willingness to Choose a Physician

A participant's intention to choose the reviewed physician was assessed with 3 items on a 7-point scale (1=would definitely not choose/definitely unwilling; 7=would definitely choose/definitely willing). The 3 items are "How likely is it that you would choose someone like Dr Smith to be your primary care doctor or surgeon?" (depending on the assigned physician condition), "How willing would you be to go to a doctor like Dr Smith for your medical care?", and "How willing would you be to recommend a doctor like Dr Smith to your family member and friends if they have the need?" An exploratory factor analysis yielded only 1 factor with an eigenvalue greater than 1, explaining 94.17% of the total variance. All 3 items have factor loadings above .90. The items were then averaged to create a composite variable (mean 4.23, SD 1.79, alpha=.97).

Manipulation Checks

Perceptions of a Physician's Technical Skills

To determine whether or not the manipulation of a physician's technical skills was successful, 7 items were used to assess participants' perceptions of this aspect (eg, *knowledgeable*, *competent*, and *skilled*). Participants were asked how well each of the 7 items described Dr Smith on a 7-point scale (1=very poorly, 7=very well). The items were averaged (mean 5.42, SD 1.34, alpha=.98). As predicted, participants assigned to conditions of high technical skills (mean 6.39, SD 0.75) perceived the physician to be more skilled technically compared with those assigned to conditions of moderate technical skills (mean 4.44, SD 1.06, $t_{483.47}$=24.78 , P<.001).

Perceptions of a Physician's Interpersonal Skills

Another set of 7 items was used to measure participants' perceptions of Dr Smith's interpersonal skills (eg, *Kind*, *Friendly*, and *Easy to talk to*) on a 7-point scale (1=very poorly, 7=very well; mean 5.26, SD 1.24, alpha=.96). As predicted,

XSL•FO
RenderX

participants assigned to conditions of high interpersonal skills (mean 6.15, SD 0.79) perceived physicians to be more skilled interpersonally compared with those assigned to conditions of moderate interpersonal skills (mean 4.38, SD 0.95, $t_{520.51}$=23.54, P<.001).

Control Variables

Current search for a physician. Participants were asked 2 questions to determine whether they were currently searching or recently intending to search for a primary care physician or back surgeon, dependent on the condition they were assigned to (eg, "How likely are you to try and find a new primary care physician or back surgeon in the next twelve months?"). They answered the questions on a 7-point scale (1=will definitely not; 7=will definitely; mean 2.88, SD 1.59).

Past Experience With a Physician

Participants were also asked 2 questions about their past experiences about looking for or having a primary care physician or back surgeon based on the physician type they were assigned to (eg, "Have you ever had a primary care physician or back surgeon?", "Have you ever searched for a primary care physician or back surgeon?"). Participants answered either yes (1) or no (2) to both questions. For participants assigned to conditions involving a primary care physician, 76.6% (209/273) reported that they have had a primary care physician and 76.2% (208/273) reported that they have searched for a primary care physician. For participants assigned to conditions involving a back surgeon, only 5.2% (14/267) reported that they have had a back surgeon and 12.4% (33/267) reported that they have searched for a back surgeon.

Perceived Reliability of Ratings

Previous research has suggested that people may perceive the reliability of Web-based ratings differently [3,35], which, in turn, may affect their willingness to choose a physician. To control for the variation, participants were asked 1 question to assess the extent to which they consider the Web-based ratings reliable (ie, "To what extent do you consider the patient ratings are reliable measures of Dr Smith's quality?"). The item was rated on a 7-point scale with the anchors 1=not reliable at all and 7=completely reliable (mean 5.14, SD 1.05).

Data Analysis

We first examined whether data met the assumption on normality. For sample sizes greater than 300, an absolute skew value greater than 2 or an absolute kurtosis greater than 7 suggests data are non-normal [36]. The dependent variable of willingness to choose a physician has a relatively normal distribution, skewness=−.22, kurtosis=−1.08. Parametric tests were used to examine the research question and hypotheses.

We then conducted a 4-way (valence of technical skills × valence of interpersonal skills × physician type × orders of ratings) analysis of covariance (ANCOVA) on people's willingness to choose a physician, controlling for current searching for a physician, past experience with a physician, and perceived reliability of Web-based ratings. As the order of ratings did not affect people's willingness to choose a physician, $F_{1,521}$=.017, P=.90, this factor was not examined further in subsequent analyses. After conducting the ANCOVA, a planned comparison t test was conducted to examine the research question on whether people place more weight on technical or interpersonal skills when selecting a physician. H1 and H2 were tested with tests of simple main effects. All analyses were run using SPSS Statistics version 25. The significance level to reject a null hypothesis was set to .05 for all analyses.

Results

Selection of a Physician in General

The research question concerns people's willingness to choose a physician with higher ratings on one aspect than the other. The ANCOVA test suggests that the 2-way interaction between ratings of technical skills and ratings of interpersonal skills significantly affected people's willingness to choose a physician, $F_{1,521}$=30.42, P<.001, η_p^2=.06. A planned comparison t test was then conducted to further examine whether people are more willing to choose a physician with higher ratings on technical skills or interpersonal skills. The condition of high ratings on both skills was assigned a weight of 2; the condition of moderate ratings on both skills was assigned a weight of −2; the condition of high ratings on technical skills and moderate ratings on interpersonal skills was assigned a weight of 1; the condition of high ratings on interpersonal skills and moderate ratings on technical skills was assigned a weight of −1. The results suggested that people were significantly more likely to choose a physician with higher ratings on technical skills than on interpersonal skills (mean 4.79, SD 1.28) compared with a physician with higher ratings on interpersonal skills than on technical skills (mean 3.06, SD 1.55, $t_{369.96}$=22.36, P<.001, Cohen d=1.22).

Importance of Technical Versus Interpersonal skills in Selection of Different Types of Physicians

The first hypothesis predicted that people had higher intention to choose a specialist who has higher ratings on technical skills than on interpersonal skills, compared with a primary care physician with the same ratings. A 3-way interaction among ratings of technical skills, ratings of interpersonal skills, and physician type was not significant, $F_{1,521}$=3.68, P=.06, η_p^2=.01. A posthoc analysis was then conducted to test the simple main effects of physician types within the interaction of technical and interpersonal skills. As predicted, participants were more willing to choose a specialist with higher ratings on technical skills than on interpersonal skills (mean 5.07, SD 1.38) compared with a primary care physician with the same ratings (mean 4.50, SD 1.36), $F_{1,521}$=5.34, P=.02. Hence, H1 was supported. Table 2 presents means and SDs of the measured variable for all conditions.

XSL•FO
RenderX

Table 2. Means and SDs of willingness to choose a physician (N=540).

Variables	Primary care physician				Specialist			
	High ratings on technical skills		Moderate ratings on technical skills		High ratings on technical skills		Moderate ratings on technical skills	
	IS[a], High	IS, Moderate	IS, High	IS, Moderate	IS, High	IS, Moderate	IS, High	IS, Moderate
Willingness to choose a physician, mean (SD)	6.07 (0.72)	4.50 (1.36)	3.09 (1.40)	3.26 (1.42)	5.91 (0.85)	5.07 (1.38)	3.06 (1.39)	2.94 (1.41)

[a]IS: interpersonal skills

The second hypothesis proposed that people were more willing to choose a primary care physician who has higher ratings on interpersonal skills than on technical skills, compared with a specialist with the same ratings. Contradictory to the prediction, the test of simple main effects suggested that people did not differ in their willingness to select a primary care physician (mean 3.09, SD 1.40) and a specialist (mean 3.06, SD 1.39) when the physician had higher ratings on interpersonal skills than on technical skills, $F_{1, 521}$=0.013, P=.91. Therefore, H2 was not supported.

Discussion

Principal Findings

Patients are increasingly empowered in this rapidly changing health care landscape. With the access to physician reviews on the Web, patients are taking a more active role in their selection of physicians. Physicians and patients have different attitudes toward reviews provided on PRWs [6]. Physicians tend to question the accuracy of Web-based reviews and view them as a threat to their reputations [4], whereas patients generally have a favorable attitude and would consult these reviews in their choice of physicians [3]. It is thus imperative to understand how Web-based reviews affect patients' perceptions and choice of physicians, which may help patients and health professionals have a better understanding of the role of PRWs in health consumerism.

Specifically, this study took the initiative to examine if Web-based physician ratings affect patients' selection of primary care physicians and specialists differently. We investigated how Web-based reviews focusing on physicians' technical and interpersonal skills affect people's intention to select different types of physicians. The results showed that people were more willing to choose a physician with higher ratings on technical skills than on interpersonal skills compared with a physician with higher ratings on interpersonal skills than on technical skills. Furthermore, people perceived technical skills as more important and were more willing to compromise on interpersonal skills in their choice of a specialist compared with a primary care physician.

This study contributes to previous research on physician selection via PRWs by experimentally testing one's preference for a physician who is high on technical skills versus interpersonal skills. Apart from previous research that relied on survey measures to assess patient's preference for a physician's skills, little work has experimentally tested the preference. By presenting patients with a mockup physician review site and a medical care vignette, we are able to aid the patients in imagining themselves in a medical situation and thus make their preferences more accessible. Before this study, it was unclear how specific factors such as rating categories or physician characteristics may affect people's choice of physicians on PRWs [14]. With an experimental design, physician types and rating categories could be separately operationalized and directly compared to examine their role in people's choice of physicians.

This study provides insight into understanding the impact of Web-based ratings on people's physician selection. Beyond valence of Web-based reviews examined in previous research [10], this study investigated how ratings of different domains could affect people's choice of physicians. The results suggested that patients tend to place more weight on technical skills than interpersonal skills when they choose physicians, regardless of physician types. Although previous research presented mixed findings on the relative importance of technical and interpersonal skills in people's physician selection [12,15,16], this study found strong support for the greater importance of technical skills over interpersonal skills. It appears that sophisticated interpersonal skills cannot make up for the lack of medical competence. Therefore, having strong interpersonal skills, although still important, does not make a physician more competitive in the health market unless the physician is also technically competent. In fact, a post hoc analysis provided further evidence by showing that people did not differ in their willingness to choose a physician with high or moderate ratings on interpersonal skills, if the physician has mediocre ratings on technical skills.

Although technical skills, in general, are more valued than interpersonal skills in patients' choice of physicians, the relative importance of these 2 skills may differ as a function of physician types [3]. It was unclear whether patients place different weightage on technical and interpersonal skills when choosing different types of physicians. To fill the gap, this study employed a controlled experiment to investigate this matter in the context of PRWs. As a primary care physician usually serves as the first check-up point before patients' visit to a specialist who diagnoses and treats more complex problems, patients tend to expect more technical skills from a specialist compared with a primary care physician. Consistent with this prediction, when people were asked to make tradeoffs between a physician's technical and interpersonal skills, they were more willing to compromise on interpersonal skills in their selection of a specialist compared with a primary care physician. Contradictory to our prediction, patients did not seem to value interpersonal skills more in their selection of primary care physicians versus specialists. Although many patients want to establish long-term relationships with their primary care physicians and value

XSL•FO
RenderX

interpersonal rapport, they may set up high standards for primary care physicians' technical skills as well. Interpersonal skills, to a certain extent, might be secondary to technical skills when people choose primary care physicians. After all, the cost of misdiagnosis or mistreatment is tremendous and may lead to irreversible consequences on patients' health. The impact of ineffective interpersonal skills on a patient's health seems to be less severe. Therefore, patients may take into account a primary care physician's interpersonal competence only if this physician meets the high standards for technical skills.

Practical Implications

These results have practical implications for physicians who have profiles on PRWs. Given that patients value technical skills over interpersonal skills, physicians who are confident with their technical skills should try to highlight this aspect in their Web presence. For instance, quite a few medical sites allow physicians to include self-descriptions or video biographies, which can serve as important venues to promote physicians' technical skills [37]. Physicians should take advantage of these channels to advocate their technical skills. Moreover, research has shown that patients and physicians tend to have different attitudes toward PRWs. Patients are generally in favor of using this service, whereas physicians have some legitimate concerns over these sites [6]. If some patients present biased opinions about a physician's technical skills, this may mislead other patients and harm the physician's reputation. To mitigate the influence of biased reviews, PRWs may consider providing both parties (ie, patients and physicians) equal opportunities to present their opinions. For example, PRWs could expand physicians' profile sections by allowing them to post multimedia contents, such as photos and videos of their work. In addition, physicians should be offered the option to respond to patient ratings and reviews on PRWs.

PRWs provide patients aggregated ratings on physicians' technical and interpersonal skills, which could be indicative of physicians' qualities and thus affect the patient choice of physicians. Although people consider ratings on both skills, ratings on technical competence, such as diagnosis and treatment, are given more weightage when choosing physicians, regardless of physician types. Therefore, PRWs could prioritize this skill set by providing more nuanced rating categories on technical skills.

Limitations

This study has several limitations that point to directions for future research. First, because previous research suggests that only a small proportion of reviews on PRWs are negative [7,38], this study did not include negative ratings. Although we deliberately excluded negative ratings to represent the reality of PRWs, it would still be worthwhile to learn how negative ratings may affect people's choice of physicians. In particular, negativity effects may take place such that patients are more impacted by negative ratings than positive ones on their selection of physicians. Under such circumstances, people may not be willing to choose a technically skilled physician who receives negative feedback on interpersonal skills.

Second, this study examined the impact of numerical ratings, but not narrative comments, on patients' willingness to choose physicians. Although patients' evaluations are primarily displayed in the format of aggregated numerical ratings on PRWs, many portals also allow patients to leave narrative comments to detail their satisfaction and dissatisfaction. Aggregated numerical ratings tend to provide patients a holistic view of physicians and the services they provide. Narrative comments, on the contrary, can capture more detailed and nuanced feedback that is not reflected in structured rating systems. A direction for future research is to investigate how numerical ratings and narrative comments work together to affect people's willingness to choose a physician, especially if 2 sources present contradictory information.

Third, this study focused on rating categories pertaining to a physician's technical and interpersonal skills. In selection of a physician, patients take into account many considerations beyond a physician's qualifications. For example, previous research has found that management practices such as punctuality and staff quality are also considered in patients' choice of physicians [31,39]. Besides reviewing a physician's qualifications, many PRWs also include rating categories on management practices. Future research should look into these aspects in addition to a physician's skills.

Fourth, despite a wide range of specialties, this study operationalized a specialist to be a back surgeon. However, it is likely that patients use different selection criteria for specialists of different types. Under certain circumstances, patients may place more weightage on a specialist's interpersonal skills than technical skills (eg, visiting a gynecologist for a check-up). Future research thus needs to examine if the influence of Web-based ratings on physician choice differs as a function of physician specialties and medical conditions.

Finally, a patient's willingness to choose a physician is influenced by a variety of factors beyond numerical ratings displayed on PRWs. For example, demographic information of a physician (eg, sex and age) and environmental factors (eg, office location) should be taken into account when examining patients' choice of physicians. Another direction for future research is to explore underlying mechanisms, especially perceptual processes, through which physician types and patient reviews affect people's choice of physicians.

Conclusions

Patients increasingly seek information on the Web when looking for health care providers. The recent growth of PRWs has resulted in efforts to investigate how these platforms affect patients' health decision making. This study sheds light on this matter by examining how Web-based ratings on a physician's technical and interpersonal skills may affect people's willingness to choose a primary care physician versus a specialist. The results suggest that patients value physicians' technical skills more than their interpersonal skills when they select physicians. Patients are more willing to make a compromise on a physician's interpersonal skills than technical skills in their choice of specialists compared with primary care physicians. Given the importance of technical skills in people's choice of physicians,

XSL•FO
RenderX

physicians who are confident in their technical skills should make efforts to promote such skills in their Web-based profiles. Carriers of PRWs should enhance the functionality of these platforms by allowing the upload of multimedia contents that physicians could use to deliver a strong Web presence. In addition, PRWs should enable physicians to respond to patients' reviews if this function is not made available on platforms. Patients, as the primary users of PRWs, need to be aware of their impact on other users and be more responsible when leaving ratings about their physicians on the Web.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - jmir_v21i6e11188_app1.pdf]

References

1. Hibbard JH, Weeks EC. Consumerism in health care. Prevalence and predictors. Med Care 1987 Nov;25(11):1019-1032. [Medline: 3320590]
2. Lorence D, Park H, Fox S. Assessing health consumerism on the Web: a demographic profile of information-seeking behaviors. J Med Syst 2006 Aug;30(4):251-258. [Medline: 16978004]
3. Hanauer D, Zheng K, Singer D, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. J Am Med Assoc 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
4. Menon AV. Do online reviews diminish physician authority? The case of cosmetic surgery in the US. Soc Sci Med 2017 Dec;181:1-8. [doi: 10.1016/j.socscimed.2017.03.046] [Medline: 28363107]
5. Gao G, McCullough J, 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 Feb 24;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]
6. Holliday A, Kachalia A, Meyer G, Sequist TD. Physician and patient views on public physician rating websites: a cross-sectional study. J Gen Intern Med 2017 Jun;32(6):626-631 [FREE Full text] [doi: 10.1007/s11606-017-3982-5] [Medline: 28150098]
7. Kadry B, Chu L, 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 Nov 16;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
8. Grabner-Kräuter S, Waiguny MK. Insights into the impact of online physician reviews on patients' decision making: randomized experiment. J Med Internet Res 2015 Apr 9;17(4):e93 [FREE Full text] [doi: 10.2196/jmir.3991] [Medline: 25862516]
9. Burkle C, Keegan MT. Popularity of internet physician rating sites and their apparent influence on patients' choices of physicians. BMC Health Serv Res 2015 Sep 26;15:416 [FREE Full text] [doi: 10.1186/s12913-015-1099-2] [Medline: 26410383]
10. Li S, Feng B, Chen M, Bell RA. Physician review websites: effects of the proportion and position of negative reviews on readers' willingness to choose the doctor. J Health Commun 2015 Apr;20(4):453-461. [doi: 10.1080/10810730.2014.977467] [Medline: 25749406]
11. Terlutter R, Bidmon S, Röttl J. Who uses physician-rating websites? Differences in sociodemographic variables, psychographic variables, and health status of users and nonusers of physician-rating websites. J Med Internet Res 2014 Mar 31;16(3):e97 [FREE Full text] [doi: 10.2196/jmir.3145] [Medline: 24686918]
12. Fung C, Elliott M, Hays R, Kahn K, Kanouse D, McGlynn E, et al. Patients' preferences for technical versus interpersonal quality when selecting a primary care physician. Health Serv Res 2005 Aug;40(4):957-977 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00395.x] [Medline: 16033487]
13. Mavis B, Vasilenko P, Schnuth R, Marshall J, Jeffs MC. Female patients' preferences related to interpersonal communications, clinical competence, and gender when selecting a physician. Acad Med 2005 Dec;80(12):1159-1165. [Medline: 16306294]
14. Li S, Lee-Won R, McKnight J. Effects of online physician reviews and physician gender on perceptions of physician skills and Primary Care Physician (PCP) selection. Health Commun 2018 May 24:1-9. [doi: 10.1080/10410236.2018.1475192] [Medline: 29792519]
15. Detz A, López A, Sarkar U. Long-term doctor-patient relationships: patient perspective from online reviews. J Med Internet Res 2013 Jul 2;15(7):e131 [FREE Full text] [doi: 10.2196/jmir.2552] [Medline: 23819959]
16. Mercado F, Mercado M, Myers N, Hewit M, Haller NA. Patient preferences in choosing a primary care physician. J Prim Care Community Health 2012 Apr 1;3(2):125-131. [doi: 10.1177/2150131911421802] [Medline: 23803456]

XSL•FO
RenderX

17. García JA, Paterniti D, Romano P, Kravitz RL. Patient preferences for physician characteristics in university-based primary care clinics. Ethn Dis 2003;13(2):259-267. [Medline: 12785424]
18. McGlone T, Butler E, McGlone VL. Factors influencing consumers' selection of a primary care physician. Health Mark Q 2002;19(3):21-37. [doi: 10.1300/J026v19n03_03] [Medline: 12077807]
19. Victoor A, Delnoij D, Friele R, Rademakers JJ. Determinants of patient choice of healthcare providers: a scoping review. BMC Health Serv Res 2012 Aug 22;12:272 [FREE Full text] [doi: 10.1186/1472-6963-12-272] [Medline: 22913549]
20. Salmon P, Young B. A new paradigm for clinical communication: critical review of literature in cancer care. Med Educ 2017 Mar;51(3):258-268 [FREE Full text] [doi: 10.1111/medu.13204] [Medline: 27995660]
21. Stiefel F, Bourquin C. Communication in oncology: now we train - but how well? Ann Oncol 2016 Dec;27(9):1660-1663. [doi: 10.1093/annonc/mdw229] [Medline: 27287204]
22. Moore S, Bopp KD. How consumers evaluate health care quality: Part I. Health Mark Q 1999;16(4):1-6. [doi: 10.1300/J026v16n04_01] [Medline: 11066713]
23. Cousin G, Mast MS, Jaunin-Stalder N. Finding the right interactional temperature: do colder patients need more warmth in physician communication style? Soc Sci Med 2013 Dec;98:18-23. [doi: 10.1016/j.socscimed.2013.08.034] [Medline: 24331877]
24. Dunlea R, Lenert L. Understanding patients' preferences for referrals to specialists for an asymptomatic condition. Med Decis Making 2015 Dec 14:0272989X14566640 [FREE Full text] [doi: 10.1177/0272989X14566640] [Medline: 25589523]
25. Hoerger T, Howard LZ. Search behavior and choice of physician in the market for prenatal care. Med Care 1995 Apr;33(4):332-349. [Medline: 7731276]
26. Cheraghi-Sohi S, Hole A, Mead N, McDonald R, Whalley D, Bower P, et al. What patients want from primary care consultations: a discrete choice experiment to identify patients' priorities. Ann Fam Med 2008;6(2):107-115 [FREE Full text] [doi: 10.1370/afm.816] [Medline: 18332402]
27. Stewart DW, Hickson GB, Pechmann C, Koslow S, Altemeier WA. Information search and decision making in the selection of family health care. J Health Care Mark 1989 Jun;9(2):29-39. [Medline: 10293672]
28. Kearley K, Freeman G, Heath A. An exploration of the value of the personal doctor-patient relationship in general practice. Br J Gen Pract 2001 Sep;51(470):712-718 [FREE Full text] [Medline: 11593831]
29. Mold J, Fryer G, Roberts AM. When do older patients change primary care physicians? J Am Board Fam Pract 2004;17(6):453-460 [FREE Full text] [Medline: 15575037]
30. Safran D, Montgomery J, Chang H, Murphy J, Rogers WH. Switching doctors: predictors of voluntary disenrollment from a primary physician's practice. J Fam Pract 2001 Feb;50(2):130-136. [Medline: 11219560]
31. Hanna N, Schoenbachler D, Gordon GL. Physician choice criteria: factors influencing patient selection of generalists versus specialists. Health Mark Q 1994;12(2):29-42. [Medline: 10141082]
32. 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]
33. Schmittdiel J, Selby J, Grumbach K, Quesenberry C. Choice of a personal physician and patient satisfaction in a health maintenance organization. J Am Med Assoc 1997 Nov 19;278(19):1596-1599. [Medline: 9370506]
34. Reimann S, Strech D. The representation of patient experience and satisfaction in physician rating sites. A criteria-based analysis of English- and German-language sites. BMC Health Serv Res 2010 Dec 7;10(17):332-341 [FREE Full text] [doi: 10.1186/1472-6963-10-332] [Medline: 21138579]
35. Rothenfluh F, Germeni E, Schulz PJ. Consumer decision-making based on review websites: are there differences between choosing a hotel and choosing a physician? J Med Internet Res 2016 Dec 16;18(6):e129 [FREE Full text] [doi: 10.2196/jmir.5580] [Medline: 27311623]
36. Kim HY. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod 2013 Feb;38(1):52-54 [FREE Full text] [doi: 10.5395/rde.2013.38.1.52] [Medline: 23495371]
37. Perrault E, Smreker KC. What can we learn from physicians' online biographies to help in choosing a doctor? Not much. A content analysis of primary care physician biographies. J Commun Healthc 2013 Nov 15;6(2):122-127. [doi: 10.1179/1753807612Y.0000000027]
38. Lagu T, Hannon N, Rothberg M, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
39. Bornstein BH, Marcus D, Cassidy W. Choosing a doctor: an exploratory study of factors influencing patients' choice of a primary care doctor. J Eval Clin Pract 2000 Aug;6(3):255-262. [Medline: 11083036]

Abbreviations

ANCOVA: analysis of covariance
PRW: physician review website

XSL•FO
RenderX

Edited by G Eysenbach; submitted 31.05.18; peer-reviewed by S Grabner-Kräuter, Y Chen, S Bidmon, M Tkáč; comments to author 30.09.18; revised version received 04.01.19; accepted 11.05.19; published 28.06.19.

Please cite as:
Li S, Hubner A
The Impact of Web-Based Ratings on Patient Choice of a Primary Care Physician Versus a Specialist: Randomized Controlled Experiment
J Med Internet Res 2019;21(6):e11188
URL: http://www.jmir.org/2019/6/e11188/
doi:10.2196/11188
PMID:31254337

©Siyue Li, Austin Hubner. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Cost-Effectiveness of Digital Health Interventions on the Management of Cardiovascular Diseases: Systematic Review

Xinchan Jiang[1*], BSc, MPhil; Wai-Kit Ming[1*], MBBS, MD, MPH, MMSCI; Joyce HS You[1*], PharmD

School of Pharmacy, The Chinese University of Hong Kong, Shatin, China (Hong Kong)

[*]all authors contributed equally

Corresponding Author:
Joyce HS You, PharmD
School of Pharmacy, The Chinese University of Hong Kong
8/F, Lo Kwee-Seong Integrated Biomedical Sciences Building, CUHK
Shatin, NT
China (Hong Kong)
Phone: 852 39436830
Fax: 852 26035295
Email: joyceyou@cuhk.edu.hk

Abstract

Background: With the advancement in information technology and mobile internet, digital health interventions (DHIs) are improving the care of cardiovascular diseases (CVDs). The impact of DHIs on cost-effective management of CVDs has been examined using the decision analytic model–based health technology assessment approach.

Objective: The aim of this study was to perform a systematic review of the decision analytic model–based studies evaluating the cost-effectiveness of DHIs on the management of CVDs.

Methods: A literature review was conducted in Medline, Embase, Cumulative Index to Nursing and Allied Health Literature Complete, PsycINFO, Scopus, Web of Science, Center for Review and Dissemination, and Institute for IEEE Xplore between 2001 and 2018. Studies were included if the following criteria were met: (1) English articles, (2) DHIs that promoted or delivered clinical interventions and had an impact on patients' cardiovascular conditions, (3) studies that were modeling works with health economic outcomes of DHIs for CVDs, (4) studies that had a comparative group for assessment, and (5) full economic evaluations including a cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, and cost-consequence analysis. The primary outcome collected was the cost-effectiveness of the DHIs, presented by incremental cost per additional quality-adjusted life year (QALY). The quality of each included study was evaluated using the Consolidated Health Economic Evaluation Reporting Standards.

Results: A total of 14 studies met the defined criteria and were included in the review. Among the included studies, heart failure (7/14, 50%) and stroke (4/14, 29%) were two of the most frequent CVDs that were managed by DHIs. A total of 9 (64%) studies were published between 2015 and 2018 and 5 (36%) published between 2011 and 2014. The time horizon was ≤1 year in 3 studies (21%), >1 year in 10 studies (71%), and 1 study (7%) did not declare the time frame. The types of devices or technologies used to deliver the health interventions were short message service (1/14, 7%), telephone support (1/14, 7%), mobile app (1/14, 7%), video conferencing system (5/14, 36%), digital transmission of physiologic data (telemonitoring; 5/14, 36%), and wearable medical device (1/14, 7%). The DHIs gained higher QALYs with cost saving in 43% (6/14) of studies and gained QALYs at a higher cost at acceptable incremental cost-effectiveness ratio (ICER) in 57% (8/14) of studies. The studies were classified as excellent (0/14, 0%), good (9/14, 64%), moderate (4/14, 29%), and low (1/14, 7%) quality.

Conclusions: This study is the first systematic review of decision analytic model–based cost-effectiveness analyses of DHIs in the management of CVDs. Most of the identified studies were published recently, and the majority of the studies were good quality cost-effectiveness analyses with an adequate duration of time frame. All the included studies found the DHIs to be cost-effective.

(J Med Internet Res 2019;21(6):e13166) doi:10.2196/13166

KEYWORDS

telemedicine; cardiovascular diseases; stroke; heart failure; myocardial infarction; heart attack; cost-effectiveness; medical economics; decision modeling; systematic review

XSL•FO
RenderX

Introduction

Digital Health Interventions

The application of information technology and mobile internet in the health care industry takes the practice of patient care to the era of digital health. Digital health is the convergence of science and technology with health, health care, living, and society [1]. According to the US Food and Drug Administration (FDA), the broad scope of digital health includes a wide range of subsectors: mobile health, telemedicine, telehealth, wearable devices, and personalized medicine [2]. Stakeholders, such as health care practitioners and researchers, have adopted digital health interventions (DHIs) aiming to promote access, reduce costs, personalize medicine, and improve outcomes of patient care. Various types of digital devices or technologies are used to deliver the health interventions, such as the short message service (SMS); mobile app; telephone; video conferencing system; digital, broadband, satellite, wireless, or Bluetooth for monitoring and transmission of physiologic data (telemonitoring [TM]); and wearable medical device [3-6].

Use of Digital Health Interventions in Cardiovascular Diseases

Cardiovascular diseases (CVDs) cause 17.9 million deaths per year, accounting for 31% of all mortality globally [7]. It is estimated that the global costs of CVDs will rise from US $863 billion in 2010 to US $1044 in 2030 [8]. The potential benefits of DHIs in CVDs were examined in clinical trials. TM was applied in an intensive follow-up of heart failure (HF) patients after discharge in the *TElemonitoring in the MAnagement of Heart Failure* study [9]. Compared with the usual care group, all-cause mortality was significantly lower in the TM group. The number of follow-up days lost to HF-related events was also significantly reduced in the TM group. A randomized controlled trial investigated the effect of CardioFit, an internet-based expert system, in patients with coronary heart disease (CHD) [10]. Patients in the CardioFit group received 5 Web-based tutorials over a period of 6 months for activity planning and tracking and were in contact with an exercise specialist. Physical activity, measured by a pedometer and self-report, was improved over a period of 12 months. A meta-analysis on 51 studies of DHIs in patients with CVDs or risk factors of CVDs reported that DHIs were associated with reduction of cardiovascular event rates and had a positive impact on risk factors for CVDs [11].

Decision Analytic Model–based Health Technology Assessment

In addition to the improved clinical outcomes of DHIs for CVD patients, evaluating the health economic outcomes is also crucial for clinicians, patients, and third-party payers in deciding the role of DHIs for CVD management. Decision analytic modeling is an approach that synthesizes cost-effectiveness evidence of health technologies and interventions in health technology assessment (HTA) [12]. This approach provides a framework to incorporate relevant clinical probabilities and cost items, simulates outcomes of disease management, and allows cost-effectiveness evaluation of medical interventions. Decision tree and Markov models are 2 commonly used forms of decision analytical modeling in health economic evaluation. In a decision tree, distinct branches are used to represent a potential set of outcomes for the patient cohort managed by an alternative treatment. Outcomes and costs for each branch are combined using branch possibilities to simulate the expected outcomes and costs for the treatment option. In a Markov model, hypothetical patients proceed through different health states over time based on transition probabilities between health states. Outcomes and costs expected by the patient cohort in an alternative treatment group are estimated from subject-time spent in various health states. The confidence level in the output of an economic modeling analysis, in relation to uncertainty in the model inputs, is typically quantified by techniques such as one-way and probabilistic sensitivity analyses. By applying model-based HTA, the cost-effectiveness impacts of various types of health technologies and interventions are compared using findings from corresponding clinical trials [12,13]. The implementation cost of DHIs is usually substantial, and HTA is, therefore, essential to inform the decision makers on the potential impact of the DHIs on both clinical and health economic outcomes [14]. The cost-effectiveness of DHIs is subject to the balance of 3 elements: clinical and economic benefits of DHI, cost of DHI, and payer's willingness-to-pay (WTP) threshold. A previous review of HTA studies on DHIs indicated that there were few health economic studies of DHIs [15].

Objective

With increasing publications on the cost-effectiveness of DHIs in CVDs, the purpose of this study is to conduct a systematic review of decision analytic model–based health economic analyses of DHIs for CVD management.

Methods

Search Strategy

The investigators developed the search strategies from September to October 2018 to include a wide range of DHIs. Analogously, search terms included different types of CVDs, such as HF, myocardial infarction, and stroke. The literature search was conducted in the following databases: Medline, Embase, Cumulative Index to Nursing and Allied Health Literature Complete, PsycINFO, Scopus, Web of Science, Center for Review and Dissemination, and Institute for IEEE Xplore. A preliminary search found an evident surge of publications on digital technologies in the field of health and medical research starting in the 2000s, and all databases were, therefore, searched back to 2001. A manual search of reference lists of both included studies and relevant systematic reviews was also conducted. Multimedia Appendix 1 provides detailed information about the search terms. This study was registered on PROSPERO with the registration number of CRD42018111473.

Inclusion and Exclusion Criteria

Full-text journal articles written in English were included if (1) the target population was patients with CVDs, (2) DHIs were aimed to promote or deliver clinical interventions and had an

XSL•FO
RenderX

impact on cardiovascular conditions, (3) decision analytic models (including decision tree and/or the Markov model) were applied to evaluate health economic outcomes of DHIs, (4) the interventions were compared with conventional care or other DHIs, and (5) a full-scale health economic evaluation was performed as a cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, or cost-consequence analysis.

The exclusion criteria included the following: (1) DHIs were only used for recording patients' information, (2) studies were conducted alongside a clinical trial, (3) quality-adjusted life years (QALYs) were not reported, or (4) the articles were reviews, protocol papers, letters, editorials, conference abstracts, poster presentations with insufficient details, or case reports.

Study Selection

After removing the duplicates, titles and abstracts were screened for eligibility. The full text of eligible articles was then reviewed for verification of eligibility. The primary search was conducted by one of the investigators (XJ). The abstracts were reviewed by two of the investigators (XJ and WM), independently. Any disagreements were discussed with the third investigator (JY) to reach a consensus. At the final stage of the full-text review, the included articles that met all the predefined criteria were read by all the investigators (XJ, WM, and JY) to confirm inclusion of the articles.

Data Extraction

A pilot data extraction was conducted by two of the investigators (XJ and WM), independently. Any discrepancy pertinent to data extraction was discussed to reach a consensus. After that, an abstraction form was adopted for guiding further data extraction. The collected information included the following items: (1) general information (including authors, title, country, and publication date), (2) study characteristics (including types of diseases and interventions), (3) methodology (including modeling method, time horizon, and perspective), and (4) summary of quantitative findings and conclusions. The primary outcomes collected were the cost-effectiveness of the DHIs, presented by incremental cost per QALY as the incremental cost-effectiveness ratio (ICER). If the ICER was not available, incremental cost and incremental QALY were assessed.

Assessment of Methodological Quality

The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist, developed by the International Society for Pharmacoeconomics and Outcomes Research for good reporting of health economic outcomes, was used to assess the methodological quality of each study [16]. The CHEERS checklist included 24 items, and the recommendations were subdivided into 6 categories: (1) title and abstract, (2) introduction, (3) methods, (4) results, (5) discussion, and (6) other. One point was assigned to each item when the quality criteria were fulfilled (and zero points for not entirely conforming to the criteria) to generate a total score (maximum score is 24). The included studies were classified into 4 quality categories: excellent (scored in 100% of the items), good quality (scored between >75% and <100% of the items), moderate quality (scored between >50% and ≤75% of the items), and low quality (scored ≤50% of the items) [17].

Two of the investigators (XJ and WM) independently assessed the quality of each study and assigned the scores based on the CHEERS checklist. Any disagreement was resolved by discussion and consensus with the third investigator (JY).

Data Analysis and Presentation

The number of studies included and excluded during the selection process was presented in a flowchart. The included studies were categorized by the type of devices or technologies used for DHI delivery. The descriptive characteristics and the study quality of the included studies and ICERs of DHIs were summarized. The DHI was categorized as cost-effective if (1) it was more effective and less costly than the comparator (DHI dominated the comparator) or (2) it was more effective at a higher cost and the ICER was less than the WTP threshold. The cost-effective DHI identified in each included study was presented.

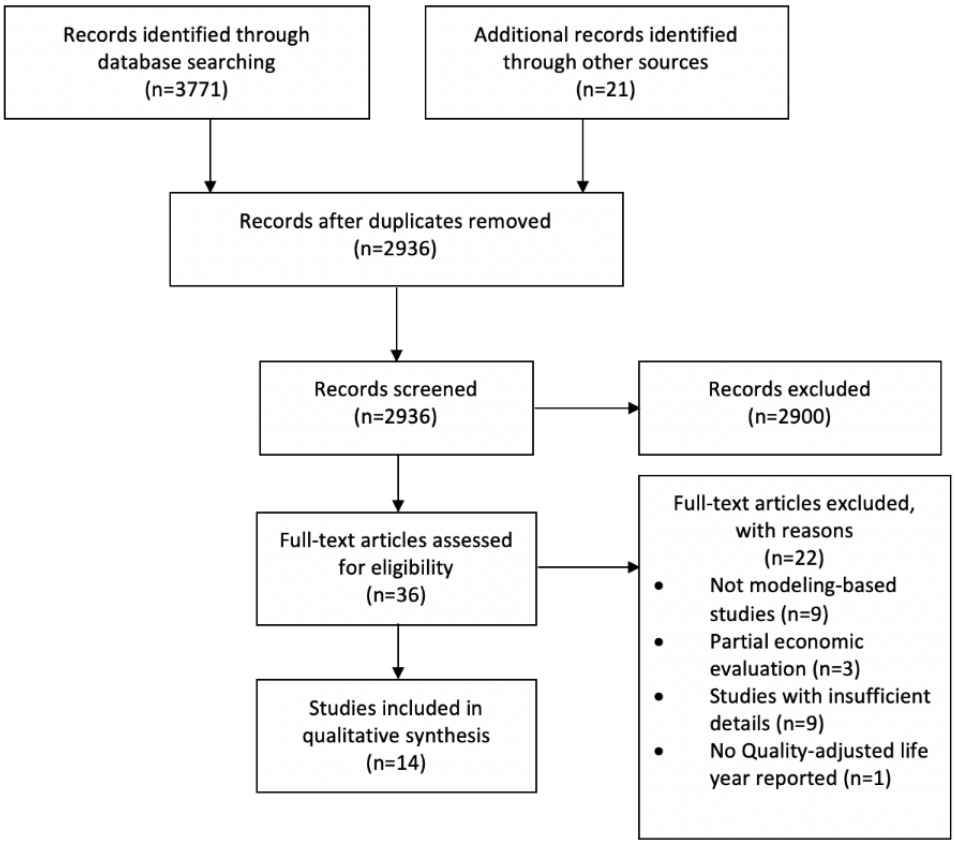
Results

Search Results

The data extraction and selected results are shown in Figure 1. The search retrieved 3771 studies from targeted databases and 21 studies from manual searches. After removal of the duplicates, 2936 articles remained. A total of 14 out of the 36 full-text articles screened according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines were included in the review [18].

XSL•FO
RenderX

Figure 1. The article selection process according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guideline.



Study Characteristics

Table 1 summarizes the characteristics and main health economic outcomes of the included studies.

A total of 10 out of the 14 studies (10/14, 71%) were conducted using the Markov model [19-21,25,27-32] and 4 studies (4/14, 29%) used the decision tree model [22-24,26]. All the 14 studies conducted (14/14, 100%) a cost-utility analysis with cost and QALYs as the outcome measures and 7 studies (7/14, 50%) also performed a cost-effectiveness analysis to include survival as the effectiveness measure [19,20,28-32]. Types of targeted CVDs were CHD [19], HF [20,21,27-31], congenital heart disease (CoHD) [22], stroke [23-26], and sudden cardiac arrest (SCA) [32]. A total of 9 studies (9/14, 64%) were published between 2015 and 2018 [19,20,23,26,28-32] and 5 studies (5/14, 36%) between 2011 and 2014 [21-22,24,25,27]. The types of devices or technologies used to deliver the health interventions were SMS (1/14, 7%) [19], telephone support (1/14, 7%) [20], mobile app (1/14, 7%) [21], video conferencing system (5/14, 36%) [22-26], digital transmission of physiologic data (TM; 5/14, 36%) [27-31], and wearable medical device (1/14, 7%) [32]. The model time horizon in 3 studies (3/14, 21%) was ≤1 year [23,24,26], whereas 10 studies (10/14, 71%) used a time horizon longer than 1 year [19,20,22,25,27-32] and 1 study (1/14, 7%) did not declare the time horizon. In total, 3 studies (3/14, 21%) received funding from the industry [19,25,30], 6 (6/14, 43%) received grants from public organizations [21,23,24,26,27,29], 1 (1/14, 7%) received no funding [22], and 4 (4/14, 29%) did not declare their funding source [20,28,31,32]. All the included studies (14/14, 100%) found the DHIs to be cost-effective: DHIs gained higher QALYs with cost-saving in 6 (6/14, 43%) studies [19,20,22,23,25,32]; DHIs gained QALYs with a higher cost at an acceptable ICER in 8 studies (8/14, 57%) [21,24,26-31]. All the 14 studies conducted a sensitivity analysis, including a probabilistic sensitivity analysis in 10 studies (10/14, 71%) [19,20,22-24,26-29,31] and a one-way sensitivity analysis in 10 studies (10/14, 71%) [22-26,28-32]. A scenario analysis, a subtype of a sensitivity analysis, was performed for the best-case, worse-case or base-case scenarios in 6 (6/14, 43%) studies to inform the optimal scenario for cost-effective use of the DHIs [19,20,26-28,31].

XSL•FO
RenderX

Table 1. General characteristics and quality assessment of the included studies.

Technologies or devices for digital health intervention delivery with references and year	Country	Targeted disease	Model type	Perspective	Time horizon	Intervention versus comparator	Incremental cost-effectiveness ratio	Cost-effective strategy (willingness-to-pay where available)	Source of funding	CHEERS[a] (%) (quality classification)
Short message service										
Burn et al, 2017 [19]	Australia	Coronary heart disease	Markov model	Australia health care system	Lifetime	TEXT ME[b] program versus UC[c]	TEXT ME program dominated UC	TEXT ME Program (Aus $64,000)	Industry	92 (Good)
Telephone support										
Grustam et al, 2018 [20]	United Kingdom	HF[d]	Markov model	UK third-party payer	0 years	TM[e] versus UC	€12,479 /QALY[f]	NTS[g] (€9000)	Not declared	92 (Good)
Grustam et al, 2018 [20]	United Kingdom	HF	Markov model	UK third-party payer	0 years	NTS versus UC	€8795 /QALY	—[h]	—	—
Grustam et al, 2018 [20]	United Kingdom	HF	Markov model	UK third-party payer	0 years	NTS versus TM	NTS dominated TM	—	—	—
Mobile apps										
Martín et al, 2014 [21]	Spain	HF	Markov model	Spain health care system	Not declared	CardioManager versus UC	€9.303/ QALY	CardioManager	Public organization	50 (Low)
Video conferencing system										
Mistry et al, 2013 [22]	United Kingdom	Congenital heart disease	Decision tree	UK health service	Lifetime	Telemedicine screening versus direct assessment	Telemedicine screening dominated direct assessment	Telemedicine screening (€20,000)	No funding	71 (Moderate)
Whetten et al, 2018 [23]	United States	Stroke	Decision tree	US health care payer	90 days	ACCESS[i] program versus no program	ACCESS program dominated no program	ACCESS program	Public organization	75 (Moderate)
Nelson et al, 2011 [24]	United States	Stroke	Decision tree	Society	90 days	Telestroke versus UC	US $108,363 /QALY	—	Public organization	92 (Good)
	United States	Stroke	Decision tree	Society	Lifetime	Telestroke versus UC	US $2449 /QALY	Telestroke (US $100,000)	—	—
Demaerschalk et al, 2013 [25]	United States	Stroke	Markov model	Society	Lifetime	Telestroke versus UC	Telestroke dominated UC	Telestroke (US $50,000)	Industry	79 (Good)
Nelson et al, 2016 [26]	United States	Stroke	Decision tree	A spoke hospital	Inpatient stay	Telestroke versus UC	US $25,991 /QALY	Telestroke (US $50,000)	Public organization	79 (Good)
Nelson et al, 2016 [26]	United States	Stroke	Decision tree	A hub hospital	Inpatient stay	Telestroke versus UC	US $47,033 /QALY	—	—	—
Telemonitoring										
Thokala et al, 2013 [27]	United Kingdom	HF	Markov model	National Health Service in England and Wales	30 years	STS HM[j] versus UC	UC dominated STS HM	TM (€20,000)	Public organization	88 (Good)

XSL•FO
RenderX

Technologies or devices for digital health intervention delivery with references and year	Country	Targeted disease	Model type	Perspective	Time horizon	Intervention versus comparator	Incremental cost-effectiveness ratio	Cost-effective strategy (willingness-to-pay where available)	Source of funding	CHEERS[a] (%) (quality classification)
Thokala et al, 2013 [27]	United Kingdom	HF	Markov model	England and Wales	30 years	TM versus UC	£11,873 /QALY	—	—	—
Thokala et al, 2013 [27]	United Kingdom	HF	Markov model	England and Wales	30 years	Structured telephone support with a human-to-human contact. versus TM	£228,035 /QALY	—	—	—
Cowie et al, 2017 [28]	United Kingdom	HF	Markov model	UK health care payer	10 years	CardioMEMS vs UC	£19,274 /QALY	CardioMEMS (US $20,000)	Not declared	79 (Good)
Sandhu et al, 2015 [29]	United States	HF	Markov model	Society	Life time	CardioMEMS versus UC	US $71,462 /QALY	CardioMEMS (US $150,000)	Public organization	88 (Good)
Schmier et al, 2017 [30]	United States	HF	Markov model	US health care payer	5 years	CardioMEMS versus UC	US $44,832 /QALY	CardioMEMS (US $100,000)	Industry	71 (Moderate)
Martinson et al, 2017 [31]	United States	HF	Markov model	US health care payer	5 years	CardioMEMS versus UC	US $12,262 /QALY	CardioMEMS (US $50,000)	Not declared	83 (Good)
Wearable medical device										
Healy et al, 2015 [32]	United States	SCA[k]	Markov model	Society	5 years	WCD[l] versus discharge home	US $26,436 /QALY	WCD (US $50,000)	Not declared	71 (Moderate)
Healy et al, 2015 [32]	United States	SCA	Markov model	Society	5 years	WCD versus SNF[m]	WCD dominated SNF	—	—	—
Healy et al, 2015 [32]	United States	SCA	Markov model	Society	5 years	WCD versus in-hospital stay	WCD dominated in-hospital stay	—	—	—

[a]CHEERS: Consolidated Health Economic Evaluation Reporting Standards is a 24-item checklist with a maximum score of 24. Studies that fulfilled 100% of the items were classified as excellent quality, those that fulfilled between >75% and <100% of the items were classified as good quality, those that fulfilled between >50% and ≤75% were classified as moderate quality, and those that fulfilled ≤50% were classified as low quality.

[b]TEXT ME: Tobacco, Exercise, and Diet Messages.

[c]UC: usual care.

[d]HF: heart failure.

[e]TM: telemonitoring.

[f]QALY: quality-adjusted life year.

[g]NTS: nurse telephone support.

[h]Not applicable.

[i]ACCESS: Access to Critical Cerebral Emergency Support Services.

[j]STS HM: structured telephone support with a human-to-machine interface.

[k]SCA: sudden cardiac arrest.

[l]WCD: wearable cardioverter-defibrillator.

[m]SNF: skilled nursing facility.

XSL•FO
RenderX

Study Quality

The percentage of items fulfilled by each study per the CHEERS checklist is shown in Table 1. The number of studies classified as excellent quality, good quality, moderate quality, and low quality were 0 (0/14, 0%), 9 (9/14, 64%), 4 (4/14, 29%), and 1 (1/14, 7%), respectively. The CHEERS checklist items and detailed quality assessment for each study are listed in Multimedia Appendix 2. In total, 3 CHEERS checklist items were fulfilled by <7 (50%) studies: (1) the abstract was fulfilled in 3 (3/14, 21%) studies [23,24,31], (2) target population and subgroups were fulfilled in 4 (4/14, 29%) studies [19,20,22,29], and (3) characterizing heterogeneity in results was fulfilled in 3 (3/14, 21%) studies [20,26,29]. In total, 7 CHEERS checklist items were fulfilled by all studies (100%): (1) title, (2) background and objectives, (3) comparators, (4) choice of health outcomes, (5) choice of model, (6) analytic methods, and (7) study findings, limitations, generalizability, and current knowledge. The remaining 14 CHEERS items were fulfilled in ≥50% of the studies: (1) setting and location [19-24,27,28,31], (2) study perspective [19,20,22,24,26-29,31,32], (3) time horizon [19,20,22-32], (4) discount rate [19,20,22,24-32], (5) measurement of effectiveness [19,20,24,25,27,31,32], (6) measurement and valuation of preference-based outcomes [19,20,23-31], (7) estimating resources and costs [19-21,23-32], (8) currency, price date, and conversion [19-21,23-32], (9) assumptions [19,20,22,24-29,31,32], (10) study parameters [19,20,22-30], (11) incremental costs and outcomes [19,20,22-32], (12) characterizing uncertainty [19,20,22-25,27-32], (13) source of funding [19,21-27,29,30], and (14) conflicts of interest [19-21,23-32].

Type of Devices or Technologies for Digital Health Intervention Delivery

Short Message Service

The Tobacco, Exercise and Diet Messages (TEXT ME) intervention sent text messages via SMS to CHD patients in addition to their usual physician counseling. The text messages included 4 types of information: general information on heart diseases, nutrition, physical activity, and smoking cessation. A total of 4 text messages were sent per week for 24 weeks. The TEXT ME program was reported to gain 1143 QALYs and save a direct medical cost of Aus $10.56 million over a lifetime horizon for a hypothetical cohort of 50,000 CHD patients in Australia [19].

Telephone Support

Structured telephone support is the use of phone calls by specialists, such as nurses, to deliver self-care support and/or management. Nurse telephone support (NTS) for HF patients, managed by a specialist nurse, included monthly assessments of symptoms, current medication, and delivery of timely feedback to physicians and patients. A cost-effectiveness analysis examined the NTS strategy for HF patients in the United Kingdom [20]. NTS gained .14 QALYs and saved €3190 (including direct medical costs) in a 20-year time horizon.

Mobile App

CardioManager, a mobile app, was divided into 3 sections to allow the patients to self-manage their heart disease conditions. An informative section that provided medical information and a patient guide. A section that recorded the user's activities (physical activities and food intake) and health measurements (vital signs). A registry of medications was also included for patients to set alarms for medication administration time. The ICER of the CardioManager was €9.303/QALY (including total direct medical costs) in Spain, yet the study did not specify the time horizon and WTP threshold [21].

Video Conferencing System

A total of 5 of the included studies evaluated the cost-effectiveness of delivering specialist consultation services via a video conferencing system for remote patients with heart disease–associated conditions [22-26].

An app of the video conferencing system, using the telemedicine equipment installed in the district hospital, allowed the remote specialist to view live or prerecorded ultrasound images of a pregnant woman and to help the local specialist to identify fetal CoHD. The telemedicine service was reported to gain .042 QALYs and save £30 (total direct medical cost) per child's lifetime in the United Kingdom [22].

The video conferencing system is applied in the Telestroke network for delivery of neurology care to remote stroke patients. Telestroke operates on a hub-and-spoke system. The spoke facilities are regional hospitals connecting to a hub hospital. The hub hospital serves as the complex stroke care provider and accepts patients from the spoke hospitals. The 4 cost-effectiveness studies of telestroke for management of acute ischemic stroke were all conducted in the United States [23-26]. The Access to Critical Cerebral Emergency Support Services with 12 partner hospitals was reported to save a total direct medical cost of US $4241 and gain .202 QALYs per patient over a period of 90 days [23]. A telestroke system with 1 hub and 8 spokes was found to be accepted as cost-effective with an ICER of US $2449/QALY [24], whereas a network with 1 hub and 7 spokes gained .022 QALYs and saved US $1436 per patient in the lifetime horizon (including both direct medical and indirect costs) [25]. A short-term analysis (duration of inpatient stay) on a network with 2 hubs and 17 spoke facilities reported the network to be accepted as cost-effective (including direct medical costs) for both the hub hospital (ICER=US$47,033/QALY) and the spoke hospital (ICER=US $25,991/QALY) [26].

Telemonitoring

TM of HF deterioration indicators (such as changes in blood pressure, intrathoracic impedance, heart rates during rest, and exertion), transmitted to health care providers for review, facilitates early detection of significant changes and allows early intervention for patients with signs of deterioration to prevent emergency admissions and avoid complications. The cost-effectiveness of TM of HF patients postdischarge was examined in the United Kingdom [27]. It was accepted to be cost-effective with an ICER of £11,873/QALY (including direct medical costs) in a 30-year time frame.

XSL•FO
RenderX

CardioMEMS, a wireless pulmonary artery pressure sensor, was the first implantable HF monitoring device approved by the FDA in 2014. The wireless sensor is permanently implanted into the distal pulmonary artery and transmits hemodynamic data to a secure website (as the patient database). Changes in pulmonary artery pressure are used, in conjunction with HF signs, to guide physicians in treatment initiation and adjustment. The cost-effectiveness of CardioMEMS was evaluated in 4 studies (3 studies in the United States and 1 study in the United Kingdom) [28-31]. CardioMEMS was accepted to be cost-effective in the 3 US studies in a lifetime horizon (ICER=US $71,462/QALY including direct and indirect medical costs) [29] and in a 5-year time frame (ICER less than US $50,000/QALY including direct medical cost) [30,31]. Similarly, a 10-year UK analysis also accepted CardioMEMS to be cost-effective with an ICER of £19,274/QALY including direct medical costs [28].

Wearable Medical Device

Wearable cardioverter-defibrillator (WCD) is an external device used for monitoring heart rhythm continuously. The WCD alarms the patient by vibration when ventricular arrhythmia (ventricular tachycardia or ventricular fibrillation) is detected. If the patient does not respond to the alarm, the WCD then delivers electric shocks to resuscitate the patient from presumed SCA. The cost-effectiveness of WCD against sudden cardiac death in patients who require temporary removal of the implantable cardioverter-defibrillator was examined [32]. WCD gained higher QALYs and saved total cost (including direct and indirect costs) when compared with discharge-to-skilled nursing facility (by .076 QALYs and US $6681) and in-hospital stay (by .039 QALYs and US $26,001).

Discussion

Principal Findings

This is the first systematic review of a decision analytic model–based cost-effectiveness analysis of DHIs in the management of CVDs. This review identified a small but growing body of evidence (14 studies) evaluating the cost-effectiveness of DHIs. These findings were similar to those of a previous review (on cost-effectiveness of telehealth interventions for HF patients) where 7 studies assessed both costs and effectiveness outcomes comprehensively [33].

The assessment of the study quality found that the majority of the methodology items (including comparators, time horizon, choice of models and health outcome measures, and analytical methods; Multimedia Appendix 2) met the requirements of the CHEERS checklist, indicating the scientific rigor of the modeling approaches applied in the included studies. All the 14 included studies were conducted in developed countries, suggesting that DHIs were more ready to be implemented in developed regions. Most of the studies were funded by a public organization (6/14, 43%), followed by the industry (3/14, 21%), showing that both the public sectors and the technology industry have a strong interest to implement cost-effective DHIs for CVD management in the health care system.

The DHIs that showed to be cost saving were SMS for CHD patients [19], telephone support for HF patients [20], wearable medical device for patients at risk of SCA [32], and video conferencing systems for prenatal screening [22] and stroke patients [23,25], whereas the DHIs found to incur higher costs were mobile app for HF patients [21], video conferencing systems for stroke management [24,26], and TM for HF patients [27-31]. The type of DHIs is, therefore, one of the influential cost drivers in the management of CVD patients.

The impact of a technology on health economics is highly subjected to the difference between the technology cost and the change (reduction or increase) in health care resource utilization as a result of the clinical effect of this technology. For instance, the effective technologies improving the survival rate of stroke patients inevitably increased the total treatment costs associated with long-term care for stroke survivors. The cost-effectiveness of the technology was influenced by both the difference in cost and in QALY gained. In this review, all studies (except one in which the time frame was not declared [21]) of DHIs for HF, CHD, CoHD, and SCA patients had used a long-term time frame (ranged from 5 years to lifelong). Both short- and long-term models were used for stroke. As shown by the studies of the Telestroke network, the cost-effectiveness of DHI improved in models with lifelong versus short-term (90 days and inpatient stay) time horizon [24-26]. The time horizon of the model needs to be adequate to capture the QALYs gained by the technology.

Limitations

This systematic review was limited by the search approach. Only studies written in the English language were included, and limited databases with fixed number keywords were used. Some relevant studies, therefore, might not be identified by this search approach, limiting both the total number and origins of studies (all in developed countries) included. The search was also limited to studies conducted by decision analytic modeling. Despite the fact that the decision analytic model–based methodology allows a cost-effectiveness analysis of multiple treatment strategies (supported by the evidence of corresponding clinical trials) for disease management, economic evaluations conducted alongside clinical trials produce valid and rigorous cost-effectiveness evidence. Further review of cost-effectiveness studies of DHIs for CVDs conducted alongside clinical trials is highly warranted.

Implications

Despite the small number of model-based health economic analyses on DHIs for CVD management, the majority were ranked to be good quality health economic evaluations and were conducted in the past 5 years. It showed an up-and-rising demand for cost-effective application of DHIs in the management of CVD patients. The development of DHIs has blossomed with the advancement of technology on the internet and mobile devices over the past 2 decades. An increasing number of digital health tools, including wearable and smart devices, make early or real-time detection, monitoring, and intervention possible for CVD patients to prevent events with high morbidity and mortality. Model-based health economic analysis is a well-accepted tool used in the technology appraisal process by national institutes, such as the National Institute for

XSL•FO
RenderX

Health and Care Excellence in the United Kingdom, to inform the clinical criteria and intervention cost for reasonable cost-effective use of health technology. In this review, the decision analytic model–based analyses evaluated the cost-effectiveness of DHIs and identified scenarios in which DHIs were likely to be accepted as cost-effective. The evidence generated by the health economic analyses facilitates the timely implementation of digital technologies in health care systems. Quality research is, therefore, highly warranted in health economic evaluation of DHIs for the management of chronic illnesses, in both developed and developing countries.

Conclusions

This is the first systematic review of decision analytic model–based cost-effectiveness analyses of DHIs in the management of CVDs. Most of the identified analyses were published recently, and the majority were good quality cost-effectiveness analyses with an adequate duration of time frame. All included studies found the DHIs to be cost-effective.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms.

[PDF File (Adobe PDF File), 157KB - jmir_v21i6e13166_app1.pdf]

Multimedia Appendix 2

CHEERS checklist and quality assessment of included studies.

[PDF File (Adobe PDF File), 425KB - jmir_v21i6e13166_app2.pdf]

References

1. Bhavnani SP, Narula J, Sengupta PP. Mobile technology and the digitization of healthcare. Eur Heart J 2016;37(18):1428-1438 [FREE Full text] [doi: 10.1093/eurheartj/ehv770] [Medline: 26873093]
2. US Food and Drug Administration. 2018. Digital Health URL: https://www.fda.gov/MedicalDevices/DigitalHealth/default.htm [accessed 2018-12-06] [WebCite Cache ID 74RziMKx0]
3. Noah B, Keller M, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. NPJ Digital Medicine 2018;1(1):p.20172 [FREE Full text]
4. World Health Orgnization. 2011. mHealth: new horizons for health through mobile technologies: second global survey on eHealth URL: https://apps.who.int/iris/handle/10665/44607 [accessed 2019-04-01] [WebCite Cache ID 77JHTdi0z]
5. Harnett B. Telemedicine systems and telecommunications. J Telemed Telecare 2006;12(1):4-15. [doi: 10.1258/135763306775321416] [Medline: 16438772]
6. Tanawuttiwat T, Garisto JD, Salow A, Glad JM, Szymkiewicz S, Saltzman HE, et al. Protection from outpatient sudden cardiac death following ICD removal using a wearable cardioverter defibrillator. Pacing Clin Electrophysiol 2014;37(5):562-568. [doi: 10.1111/pace.12319] [Medline: 24762055]
7. World Health Organization. 2018. Cardiovascular disease URL: https://www.who.int/cardiovascular_diseases/en/ [accessed 2018-12-06] [WebCite Cache ID 74S0Ny3b4]
8. Champion Advocates Programme. 2019. The costs of CVD URL: http://www.championadvocates.org/en/champion-advocates-programme/the-costs-of-cvd [accessed 2019-03-24] [WebCite Cache ID 777VWrpl9]
9. Dendale P, De Keulenaer G, Troisfontaines P, Weytjens C, Mullens W, Elegeert I, et al. Effect of a telemonitoring-facilitated collaboration between general practitioner and heart failure clinic on mortality and rehospitalization rates in severe heart failure: the TEMA-HF 1 (TElemonitoring in the MAnagement of Heart Failure) study. Eur J Heart Fail 2012;14(3):333-340 [FREE Full text] [doi: 10.1093/eurjhf/hfr144] [Medline: 22045925]
10. Reid RD, Morrin LI, Beaton LJ, Papadakis S, Kocourek J, McDonnell L, et al. Randomized trial of an internet-based computer-tailored expert system for physical activity in patients with heart disease. Eur J Prev Cardiol 2012;19(6):1357-1364. [doi: 10.1177/1741826711422988] [Medline: 21903744]
11. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. Mayo Clin Proc 2015;90(4):469-480. [doi: 10.1016/j.mayocp.2014.12.026] [Medline: 25841251]
12. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. Cost-Effectiveness in Health and Medicine Second Edition. Oxford: Oxford University Press; 2016.
13. World Health Organization. 2018. Health Technology Assessment URL: http://www.who.int/health-technology-assessment/about/Defining/en/ [accessed 2018-12-06] [WebCite Cache ID 74S0nxNEs]

XSL•FO
RenderX

14. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
15. de la Torre-Diez I, López-Coronado M, Vaca C, Aguado JS, de Castro C. Cost-utility and cost-effectiveness studies of telemedicine, electronic, and mobile health systems in the literature: a systematic review. Telemed J E Health 2015;21(2):81-85 [FREE Full text] [doi: 10.1089/tmj.2014.0053] [Medline: 25474190]
16. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 2013;16(2):231-250 [FREE Full text] [doi: 10.1016/j.jval.2013.02.002] [Medline: 23538175]
17. Geng J, Yu H, Mao Y, Zhang P, Chen Y. Cost effectiveness of dipeptidyl peptidase-4 inhibitors for type 2 diabetes. Pharmacoeconomics 2015;33(6):581-597. [doi: 10.1007/s40273-015-0266-y] [Medline: 25736235]
18. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009;151(4):264-269. [Medline: 19622511]
19. Burn E, Nghiem S, Jan S, Redfern J, Rodgers A, Thiagalingam A, et al. Cost-effectiveness of a text message programme for the prevention of recurrent cardiovascular events. Heart 2017;103(12):893-894. [doi: 10.1136/heartjnl-2016-310195] [Medline: 28235776]
20. Grustam AS, Severens JL, de Massari D, Buyukkaramikli N, Koymans R, Vrijhoef HJ. Cost-effectiveness analysis in telehealth: a comparison between home telemonitoring, nurse telephone support, and usual care in chronic heart failure management. Value Health 2018;21(7):772-782. [doi: 10.1016/j.jval.2017.11.011] [Medline: 30005749]
21. Cano Martín JA, Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Economic impact assessment from the use of a mobile app for the self-management of heart diseases by patients with heart failure in a Spanish region. J Med Syst 2014;38(9):96. [doi: 10.1007/s10916-014-0096-z] [Medline: 24994514]
22. Mistry H, Gardiner HM. The cost-effectiveness of prenatal detection for congenital heart disease using telemedicine screening. J Telemed Telecare 2013;19(4):190-196. [doi: 10.1258/jtt.2012.120418] [Medline: 23576807]
23. Whetten J, van der Goes DN, Tran H, Moffett M, Semper C, Yonas H. Cost-effectiveness of Access to Critical Cerebral Emergency Support Services (ACCESS): a neuro-emergent telemedicine consultation program. J Med Econ 2018;21(4):398-405. [doi: 10.1080/13696998.2018.1426591] [Medline: 29316820]
24. Nelson RE, Saltzman GM, Skalabrin EJ, Demaerschalk BM, Majersik JJ. The cost-effectiveness of telestroke in the treatment of acute ischemic stroke. Neurology 2011;77(17):1590-1598 [FREE Full text] [doi: 10.1212/WNL.0b013e318234332d] [Medline: 21917781]
25. Demaerschalk BM, Switzer JA, Xie J, Fan L, Villa KF, Wu EQ. Cost utility of hub-and-spoke telestroke networks from societal perspective. Am J Manag Care 2013;19(12):976-985 [FREE Full text] [Medline: 24512034]
26. Nelson RE, Okon N, Lesko AC, Majersik JJ, Bhatt A, Baraban E. The cost-effectiveness of telestroke in the Pacific Northwest region of the USA. J Telemed Telecare 2016;22(7):413-421. [doi: 10.1177/1357633X15613920] [Medline: 26541170]
27. Thokala P, Baalbaki H, Brennan A, Pandor A, Stevens JW, Gomersall T, et al. Telemonitoring after discharge from hospital with heart failure: cost-effectiveness modelling of alternative service designs. BMJ Open 2013;3(9):e003250 [FREE Full text] [doi: 10.1136/bmjopen-2013-003250] [Medline: 24048626]
28. Cowie MR, Simon M, Klein L, Thokala P. The cost-effectiveness of real-time pulmonary artery pressure monitoring in heart failure patients: a European perspective. Eur J Heart Fail 2017;19(5):661-669 [FREE Full text] [doi: 10.1002/ejhf.747] [Medline: 28176424]
29. Sandhu AT, Goldhaber-Fiebert JD, Owens DK, Turakhia MP, Kaiser DW, Heidenreich PA. Cost-effectiveness of implantable pulmonary artery pressure monitoring in chronic heart failure. JACC Heart Fail 2016;4(5):368-375 [FREE Full text] [doi: 10.1016/j.jchf.2015.12.015] [Medline: 26874380]
30. Schmier JK, Ong KL, Fonarow GC. Cost-effectiveness of remote cardiac monitoring with the CardioMEMS heart failure system. Clin Cardiol 2017;40(7):430-436 [FREE Full text] [doi: 10.1002/clc.22696] [Medline: 28272808]
31. Martinson M, Bharmi R, Dalal N, Abraham WT, Adamson PB. Pulmonary artery pressure-guided heart failure management: US cost-effectiveness analyses using the results of the CHAMPION clinical trial. Eur J Heart Fail 2017;19(5):652-660 [FREE Full text] [doi: 10.1002/ejhf.642] [Medline: 27647784]
32. Healy CA, Carrillo RG. Wearable cardioverter-defibrillator for prevention of sudden cardiac death after infected implantable cardioverter-defibrillator removal: a cost-effectiveness evaluation. Heart Rhythm 2015;12(7):1565-1573. [doi: 10.1016/j.hrthm.2015.03.061] [Medline: 25839113]
33. Grustam AS, Severens JL, van Nijnatten J, Koymans R, Vrijhoef HJ. Cost-effectiveness of telehealth interventions for chronic heart failure patients: a literature review. Int J Technol Assess Health Care 2014;30(1):59-68. [doi: 10.1017/S0266462313000779] [Medline: 24495581]

Abbreviations

CHD: coronary heart disease

XSL•FO
RenderX

CHEERS: Consolidated Health Economic Evaluation Reporting Standards
CoHD: congenital heart disease
CVD: cardiovascular disease
DHI: digital health intervention
FDA: Food and Drug Administration
HF: heart failure
HTA: health technology assessment
ICER: incremental cost-effectiveness ratio
NTS: nurse telephone support
QALY: quality-adjusted life year
SCA: sudden cardiac arrest
SMS: short message service
TEXT ME: Tobacco, Exercise and Diet Messages
TM: telemonitoring
WCD: wearable cardioverter-defibrillator
WTP: willingness-to-pay

Edited by C Eichenberg; submitted 17.12.18; peer-reviewed by Z Chen, M Lee, J Redfern, T Bergmo, J Huang, J Whetten; comments to author 06.02.19; revised version received 01.04.19; accepted 17.05.19; published 17.06.19.

Please cite as:
Jiang X, Ming WK, You JHS
The Cost-Effectiveness of Digital Health Interventions on the Management of Cardiovascular Diseases: Systematic Review
J Med Internet Res 2019;21(6):e13166
URL: http://www.jmir.org/2019/6/e13166/
doi:10.2196/13166
PMID:31210136

©Xinchan Jiang, Wai-Kit Ming, Joyce HS You. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 17.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Relationship Between Patient-Reported Outcome Measures and the Severity of Chronic Obstructive Pulmonary Disease in the Context of an Innovative Digitally Supported 24-Hour Service: Longitudinal Study

Signe Lindskrog[1], MSc; Karl Bang Christensen[1], PhD; Richard H Osborne[1,2], PhD; Søren Vingtoft[3], MD; Klaus Phanareth[3], MD, PhD; Lars Kayser[1], MD, PhD

[1]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[2]Faculty of Health, Arts and Design, Swinburne University of Technology, Hawthorn, Australia

[3]Region Zeeland, Sorø, Denmark

Corresponding Author:
Lars Kayser, MD, PhD
Department of Public Health
University of Copenhagen
Øster Farimagsgade 5
Copenhagen,
Denmark
Phone: 45 28757291
Email: lk@sund.ku.dk

Abstract

Background: Individuals with chronic obstructive pulmonary disease (COPD) live with the burden of a progressive life-threatening condition that is often accompanied by anxiety and depression. The severity of the condition is usually considered from a clinical perspective and characterized according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification of severity (1-4) and a risk assessment (A through D) that focuses on the patient's symptoms and number of exacerbations, but information about perceived health or ability to manage the condition are rarely included.

Objective: We evaluated 3 patient-reported outcome measurements (PROMs) to examine how these can be used to report on individuals with COPD who were supported by a digitally assisted intervention that aims to increase the patient's management of their condition to improve their well-being.

Methods: A total of 93 individuals with COPD were enrolled. At baseline and after 6 and 12 months, we measured self-reported self-management (Health Education Impact Questionnaire, heiQ) and health literacy (Health Literacy Questionnaire, HLQ), and physical and mental health (Short Form-36, SF-36) PROMs were collected. The scores of the 19 PROM dimensions were related to COPD severity, that is, GOLD risk assessment, pulmonary function at entry, and number of exacerbations of a period up to 12 months. The initial PROM scores were also compared with pulmonary function, exacerbations, and GOLD risk assessment to predict the number of contacts within the first 90 days.

Results: At baseline, 2 dimensions from heiQ and SF-36 Physical health differed significantly between GOLD risk factor groups, indicating more distress and poorer attitudes and health status with increasing severity (GOLD risk assessment). Pulmonary function (FEV1) was negatively associated with the severity of the condition. After 6 months, we observed an increase in heiQ6 (skill and technique acquisition) and a reduction in emotional distress. The latter effect persisted after 12 months, where heiQ4 (self-monitoring and insight) also increased. HLQ3 (actively managing my health) decreased after 6 and 12 months. The number of exacerbations and the GOLD risk factor assessment predicted the number of contacts during the first 90 days. Furthermore, 2 of the PROMS heiQ6 (skill and technique acquisition) and HLQ8 (ability to find good health information) evaluated at baseline were associated with the number of contacts within the first 90 after enrollment. The pulmonary function was not associated with the number of contacts.

Conclusions: Our data suggest that selected dimensions from HLQ, heiQ, and SF-36 can be used as PROMs in relation to COPD to provide researchers and clinicians with greater insight into how this condition affects individuals' ability to understand

XSL•FO
RenderX

and manage their condition and perception of their physical and mental health. The PROMs add to the information obtained with the clinical characteristics including the GOLD risk factor assessment.

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.6506

(J Med Internet Res 2019;21(6):e10924) doi:10.2196/10924

KEYWORDS

health literacy; empowerment; patient reported outcome measures; self-reported mental and physical health; health literacy questionnaire; health education impact questionnaire; SF-36; epital living lab; chronic obstructive pulmonary disease

Introduction

Background

People diagnosed with chronic obstructive pulmonary disease (COPD) are affected by the burden of living with a deteriorating life-threatening condition. This condition is characterized by breathlessness, with a feeling of a burden that those living with COPD need to learn how to coexist with and have control over [1]. Increasing severity of COPD may result in a decreased level of activity, experience of isolation, increasing dependency on health professionals, and development of distress and often result in comorbidities such as anxiety and depression [2]. The severity of the condition is often described according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification of severity, which builds on clinical characteristics and numbers of exacerbations, but no information about how the condition affects the patient's well-being, ability to manage their condition, or need of support is currently evaluated [3]. Here, the usage of psychometric instruments used as patient-reported outcome measures (PROM) may fill in a gap and contribute to a better understanding of how it is to live with COPD.

We here report on the utility of 3 validated PROMs to better understand people living with COPD and how a proactive and person-centered intervention, the Epital living lab, may influence the participants' ability to manage their condition and well-being.

The Epital living lab was established in 2013 to develop and test a new way to offer services to people living with COPD for increasing their independence and well-being and relieving the burden of the condition and treatment. The findings of the Epital living lab is reported elsewhere [4]. The Epital living lab was organized to provide services in a proactive way, taking full advantage of digitalization. The participants are actively involved in the management of their own condition and are supported by a 24/7 response and coordination center, which connects participants to all services they need and also initiate treatments in collaboration between the participants and care providers including medical doctors. This creates an environment where the participants are not restricted by or confined to predefined schedules of monitoring but can use the provided technologies to monitor themselves and consult health care professionals whenever they wish to.

It was hypothesized that this redesign with its involvement of the participants in taking care of their own condition would increase their health literacy, ability to manage their condition, and ultimately their physical and mental well-being.

Patient-Reported Outcome Measures

To evaluate how living with COPD affects the participants as well as the impact of the Epital living lab intervention, we used a set of multidimensional PROMs that assess diverse patient-centered outcomes—the Health Literacy Questionnaire (HLQ), Health Education Impact Questionnaire (heiQ), and the Short Form-36 (SF-36) [5-7].

The HLQ covers 9 conceptually distinct dimensions of health literacy. These dimensions reflect important elements from the perspective of the general population, health care providers, as well as policy makers. HLQ was developed using a validity-driven approach [7] to be a sensitive measure for the evaluation of interventions [7]. Recently, 2 of the 9 dimensions, HLQ6 (ability to actively engage with health care providers) and HLQ9 (ability to understand health information), were applied in 29,473 people within the Danish National Health Survey [8]. It has been validity tested in the Danish language [9] and other European languages and was found to have robust psychometric properties [10,11].

The heiQ is a widely used PROM to evaluate patient education interventions and self-management among people with a broad range of chronic conditions including COPD [12,13]. It was also developed using a validity-driven approach [6] and measures proximal outcomes related to self-management behavior across 8 dimensions. The heiQ has been found to capture dimensions strongly related to empowerment [14]. Of interest for our study, heiQ has been used in Denmark to evaluate the impact of telemedicine [15] and in Norway to access self-management in COPD [2].

The SF-36 was developed by the US Medical Outcomes Study to measure self-reported health status across 8 dimensions [16]. SF-36 is widely used and has been applied previously among patients with COPD and has identified psychological distress and poor health status when compared with the general population [17].

The aim of this study is three-fold. First, it is explored whether the clinical risk assessment of COPD is related to the individuals' perception of living with their condition as measured with the following 3 PROMs: HLQ, heiQ, and SF-36. Thereafter, it is evaluated how the Epital living lab with its innovative reorganization of services influences participants' perception of their condition measured by the 3 PROMs over a 12-month period. Finally, we explore whether clinical characteristics, commonly used by clinicians to evaluate the likelihood of deterioration or PROM scores capturing the participant's perspective, predict the participants' need of contact to the Epital response and coordination center.

XSL•FO
RenderX

Methods

Overview

This paper is part of a larger longitudinal study of the Epital living lab, which is described elsewhere [4]. In short, participants were recruited from the municipality of Lyngby-Taarbæk's rehabilitation centers via leaflets at the local pharmacies and through the Danish Lung Association. In total, 93 individuals were included from April 2013 to December 2015; 61 women (mean age 74.3 years; range 47-91 years) and 32 men (mean age 73.0 years; range 49-87 years) participated. Data about exacerbations and number of contacts were first registered by November 2013. Hence, data about these 2 parameters are not available for the first 22 participants. Each participant was offered to be part of the Epital living lab for an unlimited period, and by the end of the study period, 66 participants were still enrolled. Participants were enrolled in average for 406 days (range 8-983 days).

Patients with a COPD diagnosis were included based on the criteria that they were able to cooperate and communicate using the Epital living lab equipment and had a Mini-Mental State Examination score above 22 [4]. The test is used to examine whether the cognitive function is sufficient to cooperate when using technology. Exclusion criteria were psychosis (including severe bipolar disease) or expected lifetime less than 90 days because of a diagnosed condition other than COPD, for example, cancer.

At the time of inclusion in the study, participants received a 2-hour visit from a medical doctor in their home where they were examined to verify the diagnosis of COPD. The severity of COPD was classified based on both spirometry to measure the airflow limitation (GOLD 1-4) and according to a risk assessment combining symptoms and exacerbation (GOLD A-B-C-D) and per the criteria of the GOLD Guidelines for COPD [3].

At entry to the study, the participants were examined by the Epital eDoctors and classified based on the number of exacerbations and hospital admissions during the last year before inclusion, symptoms, and the Medical Research Council breathlessness scale and lung function.

The initial spirometry, with measurements of the forced expiratory volume during the first second (FEV1), forced vital capacity, oxygen saturation, pulse, and temperature, was supervised by the eDoctor to serve as a baseline for the algorithm developed to be used by the participants to self-monitor their condition. Full datasets were only obtained for 88 of the 93 participants. The reason was that FEV1 is a mean value of first 3 measurements, which not all participants completed. A member of the municipality's technical service team visited the participant and introduced the telemonitoring equipment: a tablet with applications for monitoring daily condition and communication with the response and coordination center by videoconference, a spirometer, a pulse oximeter, and a medical acute box [4].

Before the introduction to the equipment, the technical service team completed a mental score with the participant. The Mini Mental State Examination clarifies the mental state of the participants. If the score was above 22, the inclusion process was continued, and the participant filled in the PROMs (HeiQ, HLQ, and SF-36); if necessary, the participant was assisted by the technical service team. Filling in PROMs took approximately 30 to 50 min. If the participant appeared to be tired or not able to focus, the PROMs were collected at the next visit to avoid exhaustion. After 2 weeks, the participant was revisited by the technical service team to follow up on the usage of the equipment and to answer the participant's questions. At baseline and at 6 and 12 months, the HLQ version 1.0DK, heiQ version 2.0DK, and the SF-36 version 1.1 DK were administered.

The HLQ consists of 9 dimensions, each ranging between 4 and 6 items [7]. The response options for dimensions 1 to 5 consist of a scale of values 1 to 4 (*strongly disagree* to *strongly agree*). Dimensions 6 to 9 consist of a difficulty scale from 1 to 5 (*cannot do* to *very easy*). Each dimension score was calculated as the mean of the items comprising the dimension (19).

The heiQ consists of 8 dimensions, with 4 to 6 items with response options ranging from 1 to 4 (*strongly disagree* to *strongly agree*). Each dimension score was calculated as the mean of the items comprising the dimension [6]. Dimension 3 (emotional distress) is normally reported as impact of distress, meaning a high score reflects a high impact of distress [6]. For this study, the scale was reversed for ease of interpretation. In this way, a high score means less distress and a low score more distress.

The SF-36 includes 8 multi-item scales, comprising a total of 36 items that assess health-related dimensions. The scoring further includes 2 components' summary scores for physical and mental health [16]. To accommodate the multiple testing and reduce the risk of type 1 errors, we used only the 2 component summary scores for physical and mental health in the SF-36 questionnaire.

For all scales, missing items were imputed using the mean of other items in the dimension; however, if more than half of the items in a dimension were missing for an individual, the score for the dimension was regarded as missing.

A total of 431 questionnaires were completed, 237 (55%) under supervision and 178 (41%) without supervision (supervision status for 16 questionnaires was unregistered). Due to different inclusion periods, not all of the 93 participants completed questionnaires at 6 months and 12 months (Table 1). All questionnaires were completed between April 2013 and December 2015 except for 2 questionnaires, which were completed in January 2016. Due to a communication error with the technical service team, 15 questionnaires between January 2014 and May 2014 were not collected, resulting in a lower response rate at the 6-month follow-up.

XSL•FO
RenderX

Table 1. Questionnaire follow-up response rates.

Patient reported outcome measure	Baseline[a], n (%)	6 months[a,b], n (%)	12 months[a], n (%)
Health Literacy Questionnaire	74 (81)	36 (40)	28 (31)
Health Education Impact Questionnaire	74 (81)	36 (40)	28 (31)
Short Form 36	85 (93)	42 (46)	28 (31)

[a]Response rate.

[b]Because of a communication error with the technical service team, 15 questionnaires covering the 6-month period were not collected between January 2014 and May 2014.

From the clinical program and database used in the Epital living lab (Epiprocess), data were available on sex, age, FEV1, exacerbations, and contacts with the Epital response and coordination center. Participants' FEV1 was calculated as an average of the first 3 (supervised) spirometer tests. Percentage of expected FEV1 was based on participant's age, sex, and height. The percentage was calculated with a predicted FEV1 based on data from a Danish norm population. Sufficient data were only available for 67 of the 88 patients grouped in the GOLD risk factor groups.

For the purpose of this study, we classified the participants as having an exacerbation when they were treated for a severe exacerbation through prescription of medication (addition of oral prednisolone, broad-spectrum antibiotics, and in addition short-acting beta-2-agonist). When a participant was in contact with the response and coordination center, the Epital nurse made a note; thus, the number of notes in the records was equal to the number of contacts with the response and coordination center.

Documentation of the participants' exacerbations and contacts with the response and coordination center during the first 90 days of participation was available from November 2013 to December 2015. The participants' exacerbations and contacts with the response and coordination center was documented for each participant for the period they participated. For analysis of relation between the GOLD risk factor and PROMs to contacts and exacerbation, only the data for the first 90 days were used; 90 days was chosen as a lower number of days were assumed to not detect a sufficient number of exacerbations, and a longer period may be influenced by the variations in the participants' deterioration.

Statistical Analysis

We reported outcome measure scores, FEV1, age, contacts, and exacerbations as means with SDs. The PROMs are reported as means and percentage of maximum score and range.

One-way analysis of variance (ANOVA) was used to test for differences in PROM score, FEV1, and age between risk groups. A posteriori test (Tukey honest significant difference method) was used to determine whether differences existed between severity groups.

Changes in scores of HLQ, heiQ, and SF-36 dimensions over the 6 and 12 months were modeled using linear mixed models to account for the correlation of repeated measurements within participants.

For the analysis of association between the number of contacts within the first 90 days of participation and FEV1, GOLD risk factor assessment, exacerbations, age, sex, and PROM scores, we calculated rate ratios with corresponding 95% CIs using quasi-Poisson regression models with a scale parameter to account for overdispersion.

For all tests, a significance level of .05 was used. To accommodate for multiple testing and reduce the risk of type 1 errors, we used only the 2-component summary scores for physical and mental health in the SF-36 questionnaire. The open source statistical program R version 0.98.1028 was used for the analysis.

Ethics and Data Protection

The study was assessed and found not to need specific approval by the Regional office of the National Danish Ethics Committee (H-3-2012-FSP31). The program was also registered with the National Danish Data Agency by first the University of Copenhagen (2012-41-0384) and since January 2014 by the municipality of Lyngby-Taarbæk, Denmark (20150910229). All data were stored at the municipality and handled according to Danish legislation and regulations.

Results

Clinical Characteristics and Relationship to Health Literacy Questionnaire, Health Education Impact Questionnaire, and Short Form-36 Scores

The 93 participants recruited had a wide range of COPD severity (Table 2). There was no difference in age among the 4 risk assessment groups (A to D). The FEV1 differed across all risk assessment groups except for A, which did not differ from B and C (D and A, $P<.001$; C and B, $P<.05$; D and B, $P<.001$; D and C, $P<.01$).

In relation to the differences in PROM scores according to severity across the heiQ, HLQ, and SF-36 (19 dimensions), only 3 dimensions showed significant differences between groups (Table 2): heiQ3 (emotional distress) (D and C, $P<.05$), heiQ5 (constructive attitudes and approaches; D and C, $P<.01$) and SF-36 Physical health (D and A, $P<.001$; D and B, $P<.01$), indicating more distress, poorer attitudes, and poorer health status with increasing severity. It should be noted that heiQ3 (emotional distress) in group C was higher than the other groups, indicating less distress for this group. A similar but not significant pattern was seen in SF-36 Mental health, supporting this reduced burden of the condition in group C. FEV1, which is a direct measure of pulmonary function, declined with increasing GOLD risk assessment score.

XSL•FO
RenderX

Table 2. Baseline mean and standard deviation across chronic obstructive pulmonary disease severity.

Clinical characteristics and patient-reported outcome measure dimensions	Risk factor					Difference
	A (N=11)	B (N=23)	C (N=13)	D (N=41)	Total (N=88)	
Age (years), mean (SD)	73.1 (6.3)	72.2 (10.1)	74.8 (13.3)	74.7 (10.3)	73.9 (10.1)	—[a]
FEV1[b] (L), mean (SD)	1.47 (0.46)	1.52 (0.45)	1.15 (0.21)	0.75 (0.27)	1.11 (0.5)[c]	(A;B); (B;C); (B;D); (C;D)[d]
Expected FEV1[e], n (%)	61 (10)	61(7)	44(8)	35 (17)	45 (18)	—
HLQ1[f]: Feeling understood and supported by health care providers, mean (SD)	3.05 (0.23)	2.95 (0.63)	3.04 (0.46)	3.06 (0.54)	3.04 (0.51)	—
HLQ2: Having sufficient information to manage my health, mean (SD)	2.93 (0.37)	2.90 (0.53)	2.85 (0.36)	3.00 (0.51)	2.95 (0.47)	—
HLQ3: Actively managing my health, mean (SD)	2.73 (0.44)	2.72 (0.54)	2.78 (0.34)	2.96 (0.34)	2.84 (0.42)	—
HLQ4: Social support for health, mean (SD)	2.89 (0.59)	2.91 (0.58)	3.18 (0.42)	2.93 (0.56)	2.96 (0.54)	—
HLQ5: Appraisal of health information, mean (SD)	2.75 (0.44)	2.73 (0.63)	2.60 (0.55)	2.66 (0.47)	2.70 (0.53)	—
HLQ6: Ability to actively engage with health care providers, mean (SD)	3.98 (0.28)	3.84 (0.53)	3.75 (0.40)	3.83 (0.51)	3.85 (0.46)	—
HLQ7: Navigating the health care system, mean (SD)	3.73 (0.56)	3.59 (0.50)	3.47 (0.35)	3.66 (0.56)	3.62 (0.50)	—
HLQ8: Ability to find good health information, mean (SD)	3.87 (0.49)	3.82 (0.59)	3.58 (0.37)	3.74 (0.40)	3.75 (0.46)	—
HLQ9: Understanding health information well enough to know what to do, mean (SD)	3.82 (0.36)	3.85 (0.52)	3.80 (0.27)	3.91 (0.41)	3.87 (0.41)	—
heiQ1[g]: Health directed activities, mean (SD)	3.09 (0.56)	2.85 (0.69)	3.15 (0.74)	2.73 (0.77)	2.89 (0.72)	—
heiQ2: Positive and active engagement in life, mean (SD)	2.98 (0.55)	3.18 (0.57)	3.13 (0.42)	2.99 (0.43)	3.07 (0.48)	—
heiQ3: Emotional distress, mean (SD)	2.74 (0.63)	2.67 (0.72)	3.07 (0.68)	2.34 (0.68)	2.61 (0.71)[h]	(C;D)[i]
heiQ4: Self-monitoring and insight, mean (SD)	3.03 (0.32)	3.02 (0.42)	3.14 (0.49)	3.12 (0.41)	3.08 (0.40)	—
heiQ5: Constructive attitudes and approaches, mean (SD)	3.13 (0.46)	3.15 (0.50)	3.33 (0.38)	2.78 (0.50)	3.02 (0.51)[j]	(C;D)[k]
heiQ6: Skill and technique acquisition, mean (SD)	3.11 (0.39)	2.93 (0.43)	3.00 (0.52)	2.95 (0.41)	2.98 (0.42)	—
heiQ7: Social integration and support, mean (SD)	3.05 (0.43)	3.03 (0.56)	3.07 (0.52)	3.03 (0.54)	3.04 (0.51)	—
heiQ8: Health services navigation, mean (SD)	3.24 (0.39)	3.28 (0.41)	3.40 (0.44)	3.23 (0.54)	3.28 (0.47)	—
SF-36: Physical, mean (SD)	45.9 (8.1)	40.5 (10.8)	39.3 (9.7)	31.7 (7.5)	37.2 (10.1)[b]	(A;D) (B;D)[l]
SF 36: Mental, mean (SD)	49.5 (14.9)	46.5 (16.0)	53.3 (14.1)	40.8 (13.3)	45.8 (10.1)	—

[a]No difference.

[b]FEV1: forced expiratory volume during the first second.

[c]$P<.001$.

[d]Regarding FEV1, all GOLD risk factor groups differ from each other except for A, which did not differ from B and C. D and A, $P<.001$; C and B, $P<.05$; D and B, $P<.001$; D and C, $P<.01$.

[e]Number in GOLD risk factor groups: A=9, B=15, C=9, and D=34.

[f]HLQ: Health Literacy Questionnaire.

[g]heiQ: Health Education Impact Questionnaire.

[h]$P<.05$.

[i]The difference is between C and D, $P<.05$.

[j]$P<.01$.

[k]The difference is between C and D, $P<.01$

[l]The difference is between A and D, $P<.001$; D and B, $P<.01$

XSL•FO
RenderX

Using correlation test, FEV1 was found to be positively associated with heiQ2 (positive and active engagement in life, $P<.05$), heiQ3 (emotional distress, $P<.05$), heiQ5 (constructive attitudes and approaches, $P<.001$), and heiQ8 (health service navigation, $P<.05$) but also with SF-36 Physical health ($P<.001$). Furthermore, we found a negative association between FEV1 and HLQ3 (actively managing my health, $P<.05$).

Over a 12-month period, various changes in PROMs were found. After 6 months, an increase was observed for heiQ6 (skill and technique acquisition) as well as in the reversed heiQ3 (emotional distress). The latter effect persisted after 12 months. HeiQ4 (self-monitoring and insight) also increased slightly. Meanwhile, HLQ3 (actively managing my health) decreased after 6 and 12 months. No changes in the SF-36 mental and physical components were observed (Table 3).

Table 3. Change in patient-reported outcome measures over 6 and 12 months using linear mixed models to account for the correlation of repeated measurements within participants.

Patient-reported outcome measures	Baseline (N=74)[a], mean	6 months (N=36)[b]		12 months (N=28)	
		Mean	Mean baseline to 6-month difference (95% CI)	Mean	Mean baseline to 12-month difference (95% CI)
HLQ1[c]: Feeling understood and supported by health care providers[d]	3.0	3.0	−0.04 (−0.22 to 0.13)	3.0	0.00 (−0.19 to 0.19)
HLQ2: Having sufficient information to manage my health[d]	2.9	3.0	0.03 (−0.12 to 0.17)	2.9	0.00 (−0.15 to 0.16)
HLQ3: Actively managing my health[d]	2.8	2.8	−0.06 (−0.16 to 0.05)	2.7	−0.14[e] (−0.25 to −0.02)
HLQ4: Social support for health[d]	3.0	3.0	0.04 (−0.11 to 0.18)	2.9	−0.06 (−0.22 to 0.09)
HLQ5: Appraisal of health information[d]	2.7	2.8	0.06 (−0.08 to 0.21)	2.6	−0.08 (−0.23 to 0.08)
HLQ6: Ability to actively engage with health care providers[f]	3.8	3.9	0.10 (−0.07 to 0.27)	3.7	−0.07 (−0.26 to 0.12)
HLQ7: Navigating the health care system[f]	3.6	3.6	0.01 (−0.17 to 0.18)	3.6	−0.04 (−0.24 to 0.14)
HLQ8: Ability to find good health information[f]	3.7	3.7	−0.04 (−0.23 to 0.13)	3.6	−0.10 (−0.29 to 0.10)
HLQ9: Understanding health information well enough to know what to do[f]	3.9	3.9	0.00 (−0.16 to 0.16)	3.9	−0.05 (−0.22 to 0.13)
heiQ1[g]: Health directed activities[d]	2.9	2.9	−0.01 (−0.21 to 0.19)	2.8	0.06 (−0.17 to 0.28)
heiQ2: Positive and active engagement in life[d]	3.1	3.1	−0.03 (−0.17 to 0.10)	3.1	0.05 (−0.10 to 0.20)
heiQ3: Emotional distress[d,h]	2.6	2.8	0.24[i] (0.08 to 0.40)	2.8	0.20[e] (0.02 to 0.38)
heiQ4: Self-monitoring and insight[d]	3.0	3.0	0.01 (−0.14 to 0.15)	3.2	0.16[e] (0.00 to 0.32)
heiQ5: Constructive attitudes and approaches[d]	3.0	3.1	0.08 (−0.07 to 0.23)	3.1	0.06 (−0.10 to 0.23)
heiQ6: Skill and technique acquisition[d]	2.9	3.1	0.17[e] (0.02 to 0.33)	3.0	0.14 (−0.04 to 0.32)
heiQ7: Social integration and support[d]	3.0	3.0	−0.05 (−0.18 to 0.07)	3.0	−0.03 (−0.17 to 0.11)
heiQ8: Health services navigation[d]	3.2	3.3	0.05 (−0.10 to 0.21)	3.2	0.02 (−0.15 to 0.20)
SF-36[j]: Physical	37.2	37.2	−0.01 (−2.5 to 2.5)	36.1	−1.1 (−4.0 to 1.8)
SF-36: Mental	45.5	47.3	1.8 (−1.8 to 5.4)	47.0	1.5 (−2.6 to 5.7)

[a]N=82 for SF-36 Physical SF-36 Mental.

[b]N=42 for SF-36 Physical SF-36 Mental.

[c]HLQ: Health Literacy Questionnaire.

[d]Range 1-4.

[e]$P<.05$.

[f]Range 1-5.

[g]heiQ: Health Education Impact Questionnaire.

[h]Scale reversed, that is, high score is low emotional distress.

[i]$P<.01$.

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

XSL•FO
RenderX

Association Between Patient-Reported Outcome Measures and the Number of Contacts to the Response and Coordination Center

As seen in Table 4, the number of contacts and exacerbations increased with increasing COPD severity. There was a positive association between severity and contacts (P<.05) and severity and exacerbations (P<.001), respectively.

To determine whether the number of contacts is associated with clinical characteristics or PROM scores at the baseline, a quasi-Poisson regression model was constructed (Table 5).

Of the clinical characteristics, the number of exacerbations and severity expressed as the GOLD risk factor assessment predicted the number of contacts during the 90 days.

There was no association between baseline-FEV1, sex, or age in relation to the number of contacts.

Of the PROMs, we found that an increase in HLQ3 (actively managing my health) predicted higher number of contacts (P<.05); however, adjustment for severity proved this effect insignificant. In contrast, SF-36 Physical health component score predicted a decrease in the number of contacts, but it was also found to be insignificant after adjustment for severity of COPD. Higher baseline scores on heiQ6 (skill and technique acquisition) and HLQ8 (ability to find good health information) predicted lower number of contacts. The association between number of contacts and heiQ6 was significant both with and without adjustment for severity, whereas the association with HLQ8 was only significant after adjustment.

There were no associations between PROM scores at entry and exacerbations over the 90 days (data not shown).

Table 4. Number of contacts with the response and coordination center and registered severe exacerbations grouped by chronic obstructive pulmonary disease severity over 90 days.

Number	A	B	C	D	Total
Contacts, mean (SD)	6.0 (2.8)	7.6 (6.6)	12.0 (20.1)	15.9 (16.0)	11.6 (13.8)
Exacerbations, mean (SD)	0.0 (0.0)	0.21 (0.42)	0.42 (1.13)	0.54 (0.83)	0.36 (0.72)

XSL•FO
RenderX

Table 5. Association between clinical characteristics, patient-reported outcome measures, and number of contacts during the first 90 days after entry (rate ratios and 95% CIs). Column 2 presents the unadjusted rate ratio and column 3 gives rate ratio adjusted for GOLD severity. Rate ratio greater than 1 indicates a higher frequency and lesser than 1 a lower frequency in number of contacts with increased values of the variable.

Variable	Baseline prediction of number of future contacts over 90 days, rate ratio (95% CI)	
	Unadjusted	Adjusted for GOLD[a] severity score
FEV1[b]	0.76 (0.42-1.32)	1.96 (0.92-4.02)
GOLD severity score	0.71[c] (0.54-0.90)	NA[d] [variable used as covariate]
Sex (1=men, 0=female)	1.04 (0.56-1.85)	1.13 (0.64-1.93)
Age (years)	0.99 (0.96-1.02)	0.99 (0.96-1.01)
Exacerbations (number)	1.99[e] (1.64-2.40)	1.89[e] (1.54-2.31)
HLQ1[f]: Feeling understood and supported by health care providers	1.14 (0.63-2.14)	0.97 (0.57-1.70)
HLQ2: Having sufficient information to manage my health	1.18 (0.63-2.21)	0.93 (0.52-1.68)
HLQ3: Actively managing my health	1.94[c] (1.02-3.71)	1.47 (0.75-2.86)
HLQ4: Social support for health	1.02 (0.58-1.76)	0.92 (0.54-1.55)
HLQ5: Appraisal of health information	1.24 (0.70-2.17)	1.17 (0.70-1.94)
HLQ6: Ability to actively engage with health care providers	0.85 (0.46-1.61)	0.87 (0.50-1.52)
HLQ7: Navigating the health care system	0.88 (0.52-1.53)	0.81 (0.51-1.34)
HLQ8: Ability to find good health information	0.55 (0.30-1.00)	0.55[c] (0.32-0.96)
HLQ9: Understanding health information well enough to know what to do	0.80 (0.41-1.57)	0.72 (0.39-1.33)
heiQ1[g]: Health directed activities	0.72 (0.47-1.08)	0.78 (0.54-1.13)
heiQ2: Positive and active engagement in life	0.92 (0.52-1.62)	0.94 (0.54-1.62)
heiQ3: Emotional distress[h]	0.81 (0.56-1.18)	0.90 (0.63-1.28)
heiQ4: Self-monitoring and insight	1.35 (0.68-2.60)	1.18 (0.62-2.20)
heiQ5: Constructive attitudes and approaches	0.62 (0.36-1.07)	0.74 (0.43-1.26)
heiQ6: Skill and technique acquisition	0.47[c] (0.26-0.86)	0.49[c] (0.29-0.84)
heiQ7: Social integration and support	0.76 (0.43-1.34)	0.73 (0.43-1.23)
heiQ8: Health services navigation	0.64 (0.34-1.18)	0.67 (0.40-1.14)
SF-36[i]: Physical	0.96[j] (0.94-0.99)	0.98 (0.95-1.00)
SF 36: Mental	0.98 (0.96-1.01)	0.99 (0.97-1.01)

[a]GOLD: Global Initiative for Chronic Obstructive Lung Disease.

[b]FEV1: forced expiratory volume during the first second.

[c]*P*<.05.

[d]N/A: not applicable.

[e]*P*<.001.

[f]HLQ: Health Literacy Questionnaire.

[g]heiQ: Health Education Impact Questionnaire.

[h]Scale reversed, that is, high score is low emotional distress.

[i]SF-36. Short Form-36.

[j]*P*<.01.

Discussion

In this study, we have systematically explored, prospectively, the utility of 3 PROMs to evaluate the impact of living with COPD and how an innovative reorganization of services improves participants' understanding, self-management, and well-being. We used 3 psychometrically robust PROMs that have previously been used to assess health literacy, empowerment, and self-management in COPD and self-reported physical and mental health.

XSL•FO
RenderX

When this study was initiated in 2013, there was limited literature on the HLQ and heiQ in relation to chronic conditions, such as COPD. Given that the Epital living lab was a substantial clinical and self-care change for patients, we regarded it as critical to explore the participants' severity of COPD in relation to their reported impact of living with the condition and how this changed over time; hence, we employed the HLQ, heiQ, and SF-36 to assist with this task.

Characteristics of Participants in Relation to Severity of COPD

The increase in skills and behavior related to being more active in self-managing while experiencing a highly active disease is likely to reflect direct learning over time from closer contact and regular interactions with the health care services and through increasing competence in engaging with the technology. Our data show that those who had increased contact with the 24/7 response and coordination center tended to report a decrease in heiQ3 (emotional distress) as well as a higher heiQ6 (skill and technique acquisition) at entry to the study.

Overall, there was a tendency for those in risk assessment group C, the second highest category, to report a lower burden related to COPD compared with the other groups. This was evident through heiQ3 (emotional distress) and SF-36 Mental health scores. This may be explained by the fact that COPD patients in group C, and also D, were followed up by hospital specialists in Denmark, in contrast to patients in groups A and B, who were followed up by their general practitioners. The poorer mental well-being of group D may reflect the very poor and debilitating health status they experience that is not alleviated by specialist contacts. This is only speculative and needs to be confirmed in further studies, examining a larger sample and also applying a mixed-methods approach. Recently, a Norwegian study examined the relation between heiQ domains, burden of the condition, and severity in people with COPD. In agreement with our findings, they demonstrated high distress but in contrast to our study, also showed that all 8 heiQ dimensions, except for the heiQ4 (self-monitoring and insight), were directly associated with the burden of living with COPD [2].

The Impact of the Intervention Over a 6- and 12-Month Period

The 24/7 access to health resources provided by the Epital living lab was associated with a reduction in the burden of the condition (lower emotional distress—ie, reduced negative affective responses to the disease). This may be due to their knowledge that assistance and treatment options are always only one click away. The training during and after the inclusion and the regular use of the technology may also have resulted in an increase in both heiQ4 (self-monitoring and insight) and heiQ6 (skill and technique acquisition) as the participants acquire confidence in using the technical solution with the tablet, apps, pulse oximeter, and spirometer. Surprisingly, HLQ3 (actively managing my health) declined over the period. This may be related to the intervention being on hand 24/7. It is possible that some reduced vigilance in self-care may have occurred given the constant presence of strong and immediate support and security.

The finding that heiQ3 (emotional distress) and heiQ4 (self-monitoring and insight) increase over 12 months are in line with a regional study in Denmark with telemedicine interventions regarding COPD where 3 heiQ dimensions were found to increase over a 6-month period: heiQ3 (emotional distress), heiQ4 (self-monitoring and insight), and heiQ8 (health services navigation) [15], with the first 2 being in accordance with our results. Interestingly, we also found an increase in heiQ6 (skill and technique acquisition) but not in heiQ8 (health services navigation). This may be explained by differences in the design between the studies. In the Epital living lab, participants had continuous access to the services and were supported by municipality nurses, who were in turn, supported by eDoctors. The other study did not offer daily access and was administered by a hospital. These findings indicate that digitally supported health services may decrease the distress and relieve the burden of COPD and contribute to better self-management, self-monitoring, and insight.

Predictors of Number of Contacts During the First 90 Days

As anticipated, the severity of the condition as estimated by the GOLD risk assessment groups predicted the number of contacts with the response and coordination center during the first 90 days.

This may be the result of the high number of exacerbations found in the higher risk groups. This supports the GOLD risk factor assessment as a useful tool to predict necessary resources [3]. Interestingly, the participants' FEV1 at entry did not correlate to the number of contacts. This supports the intent of the new GOLD risk assessment, as FEV1 is not an essential indicator for the need of contacts.

Only 2 of the PROMs, heiQ6 (skill and technique acquisition) and HLQ8 (ability to find good health information), were independently associated with the number of contacts after a correction for severity of the condition. Higher scores for these 2 domains were related to a lower number of contacts, which suggests that development of skills to handle the condition and ability to find information reduces the need for assistance from the response and coordination center. It may be of value to focus on these competences for people living with COPD to reduce the burden of disease they experience.

Interestingly, our findings suggest that selected PROMs can contribute to a richer understanding of how participants are influenced by their condition. An example is how the HLQ3 (actively managing my health) is higher in GOLD group D, indicating that a higher number of contacts to the health professionals and over time decreases when participating in the Epital living lab intervention. This indicates that this PROM may capture a new perspective on the burden of the condition.

Overall, our results demonstrate that selected dimensions from heiQ and HLQ may be used to better understand people living with COPD and how a digitally assisted intervention affects them with less emotional distress, acquired skills to handle the condition, and at the same time, an indication of feeling less actively involved. Whether the latter is due to a supportive environment or relief of the burden of the condition remains to

XSL•FO
RenderX

be further investigated in larger studies supplemented with inclusion of interviews and observations. In pursuing the idea of using selected dimensions from heiQ and HLQ, we developed a new instrument inspired by our findings from this study and other reports, where we combine heiQ dimensions 3 to 6 and HLQ dimensions 1 and 4 with the 7 dimensions of the recently developed eHealth Literacy Questionnaire (eHLQ), resulting in an instrument to measure readiness and enablement of technology called Readiness and Enablement Index for Health Technology [18].

Limitations of the Study

The investigators expected that the participants in the living lab would develop their knowledge, skills, attitudes, and motivation during the study. The relatively few dimensions, which increased and the decreased in HLQ3 (actively managing my health), was unexpected. It is likely that we in the implementation of the Epital living lab as an action research project focused too much on the technical setup, stability and usability of solutions, and the medical support when needed. We therefore recommend that when the Epital living lab is implemented at other sites, more attention should be paid to the development of activities strengthening the participants' understanding of how to live with COPD and other conditions. Examples are educational and training programs, coaching, and other tools such as the guided self-determination program [19,20], which is designed to increase life skills and empower the participants.

Another limitation of the study was that the number of participants was small, and it was not possible to design a control study or match with controls in other studies. There is a risk of type 2 error because of the limited number of participants and the further reduction of participants after 6 and 12 months.

The convenience sampling in a community with a relatively high average income is also a weakness, limiting generalization to other settings. However, given the aims of the study, we expect the data are internally valid and provide reasonable guidance regarding selection of PROMs for wider study of the Epital living lab.

A final weakness is that we were not able to include some of the most recent PROMs that measure eHealth literacy at the time of the data collection. A recently published tool, the eHLQ, may have provided clearer indicators of participants' ability to engage with the technology and reasons for limited participation and consequent well-being improvement [21].

Conclusions

In this study, we demonstrated how the HLQ, heiQ, and SF-36 can be used as PROMs in relation to COPD to provide researchers and clinicians with greater insight into how this condition affects individuals' ability to understand and manage their condition and how they perceive their physical and mental health. The PROMs add to the information obtained with the clinical characteristics including the GOLD risk assessment.

At baseline, the PROM data from the GOLD risk group D provided a more nuanced picture of how living with COPD affects their well-being. The combination of self-reported increased emotional distress, lower self-reported physical health, and a less constructive approach and attitude to their condition provides a clear impression of how frail they are.

The PROMs also help to understand how the Epital living lab affects the participants over a 12-month period, including the potential for increasing skills and at the same time reducing both emotional distress and the perception of being active in managing their condition.

Acknowledgments

RHO was funded through a National Health and Medical Research Council Senior Research Fellowship #APP1059122. KP and SV have received support from the 3P project from the Norwegian health foundation, in the period 2016-2018 while contributing to the manuscript.

The authors wish to thank Emily Duminski for assisting with preparing the manuscript. The authors thank the health professionals, patients, administrators, and politicians at the municipality of Lyngby-Taarbæk for their active participation in the project. They would also like to thank Charlotte Dorph Lyng and Jakob Sylvest Nielsen for their active engagement.

Conflicts of Interest

None declared.

References

1. Cooney A, Mee L, Casey D, Murphy K, Kirwan C, Burke E, PRINCE Team. Life with chronic obstructive pulmonary disease: striving for 'controlled co-existence'. J Clin Nurs 2013 Apr;22(7-8):986-995. [doi: 10.1111/j.1365-2702.2012.04285.x] [Medline: 23279604]
2. Bringsvor HB, Skaug K, Langeland E, Oftedal BF, Assmus J, Gundersen D, et al. Symptom burden and self-management in persons with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2018 Jan;13:365-373 [FREE Full text] [doi: 10.2147/COPD.S151428] [Medline: 29416327]
3. Global Initiative for Chronic Obstructive Lung Disease. 2018. 2018 Global Strategy for Prevention, Diagnosis and Management of COPD URL: https://goldcopd.org/archived-reports/ [accessed 2019-05-10] [WebCite Cache ID 78GKGtX5H]

XSL•FO
RenderX

4. Phanareth K, Vingtoft S, Christensen A, Nielsen J, Svenstrup J, Berntsen G, et al. The epital care model: a new person-centered model of technology-enabled integrated care for people with long term conditions. JMIR Res Protoc 2017 Jan 16;6(1):e6 [FREE Full text] [doi: 10.2196/resprot.6506] [Medline: 28093379]
5. Ruta DA, Garratt AM, Leng M, Russell IT, MacDonald LM. A new approach to the measurement of quality of life. The Patient-Generated Index. Med Care 1994 Nov;32(11):1109-1126. [Medline: 7967852]
6. Osborne RH, Elsworth GR, Whitfield K. The Health Education Impact Questionnaire (heiQ): an outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. Patient Educ Couns 2007 May;66(2):192-201. [doi: 10.1016/j.pec.2006.12.002] [Medline: 17320338]
7. Osborne RH, Batterham RW, Elsworth GR, Hawkins M, Buchbinder R. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013;13:658 [FREE Full text] [doi: 10.1186/1471-2458-13-658] [Medline: 23855504]
8. Friis K, Lasgaard M, Osborne RH, Maindal HT. Gaps in understanding health and engagement with healthcare providers across common long-term conditions: a population survey of health literacy in 29,473 Danish citizens. BMJ Open 2016 Jan 14;6(1):e009627 [FREE Full text] [doi: 10.1136/bmjopen-2015-009627] [Medline: 26769783]
9. Maindal H, Kayser L, Norgaard O, Bo A, Elsworth G, Osborne R. Cultural adaptation and validation of the Health Literacy Questionnaire (HLQ): robust nine-dimension Danish language confirmatory factor model. Springerplus 2016;5(1):1232 [FREE Full text] [doi: 10.1186/s40064-016-2887-9] [Medline: 27536516]
10. Nolte S, Osborne RH, Dwinger S, Elsworth GR, Conrad ML, Rose M, et al. German translation, cultural adaptation, and validation of the Health Literacy Questionnaire (HLQ). PLoS One 2017;12(2):e0172340 [FREE Full text] [doi: 10.1371/journal.pone.0172340] [Medline: 28234987]
11. Kolarcik P, Cepova E, Madarasova GA, Elsworth GR, Batterham RW, Osborne RH. Structural properties and psychometric improvements of the Health Literacy Questionnaire in a Slovak population. Int J Public Health 2017 Jun;62(5):591-604. [doi: 10.1007/s00038-017-0945-x] [Medline: 28258403]
12. Bélanger A, Hudon C, Fortin M, Amirall J, Bouhali T, Chouinard M. Validation of a French-language version of the health education impact Questionnaire (heiQ) among chronic disease patients seen in primary care: a cross-sectional study. Health Qual Life Outcomes 2015 May 24;13:64 [FREE Full text] [doi: 10.1186/s12955-015-0254-0] [Medline: 26306793]
13. Nolte S, Elsworth GR, Sinclair AJ, Osborne RH. The extent and breadth of benefits from participating in chronic disease self-management courses: a national patient-reported outcomes survey. Patient Educ Couns 2007 Mar;65(3):351-360. [doi: 10.1016/j.pec.2006.08.016] [Medline: 17027221]
14. Barr P, Scholl I, Bravo P, Faber M, Elwyn G, McAllister M. Assessment of patient empowerment--a systematic review of measures. PLoS One 2015;10(5):e0126553 [FREE Full text] [doi: 10.1371/journal.pone.0126553] [Medline: 25970618]
15. Lee A, Sandvei M, Hosbond T, Petersen J, Kristiansen K. MedCom with Region Midtjylland and Region Hovedstaden. 2015. Klinisk Integreret Hjemmemonitorering (KIH): Slutrapportering til Fonden for Velfærdsteknologi URL: https://www.digst.dk/Digital-velfaerd/Rapporter-og-analyser/Sundhed/Klinisk-integreret-hjemmemonitorering [accessed 2019-05-10] [WebCite Cache ID 78GJFsIoa]
16. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
17. Lu Y, Nyunt M, Gwee X, Feng L, Feng L, Kua E, et al. Life event stress and chronic obstructive pulmonary disease (COPD): associations with mental well-being and quality of life in a population-based study. BMJ Open 2012;2(6):a [FREE Full text] [doi: 10.1136/bmjopen-2012-001674] [Medline: 23166130]
18. Kayser L, Rossen S, Karnoe A, Elsworth G, Vibe-Petersen J, Christensen JF, et al. Development of the multidimensional Readiness and Enablement Index for Health Technology (READHY) tool to measure individuals' health technology readiness: initial testing in a cancer rehabilitation setting. J Med Internet Res 2019 Feb 12;21(2):e10377 [FREE Full text] [doi: 10.2196/10377] [Medline: 30747717]
19. Zoffmann V, Lauritzen T. Guided self-determination improves life skills with type 1 diabetes and A1C in randomized controlled trial. Patient Educ Couns 2006 Dec;64(1-3):78-86. [doi: 10.1016/j.pec.2005.11.017] [Medline: 16720089]
20. Rasmussen B, Currey J, Haigh B, Dunning T, Zoffmann V. An online self-management intervention for young adults with type 1 diabetes: Guided Self-Determination program. Australian Diabetes Educator 2017 Sep 01;20(3):19 [FREE Full text]
21. Kayser L, Karnoe A, Furstrand D, Batterham R, Christensen K, Elsworth G, et al. A multidimensional tool based on the eHealth Literacy Framework: development and initial validity testing of the eHealth literacy Questionnaire (eHLQ). J Med Internet Res 2018 Dec 12;20(2):e36 [FREE Full text] [doi: 10.2196/jmir.8371] [Medline: 29434011]

Abbreviations

ANOVA: analysis of variance
COPD: chronic obstructive pulmonary disease
eHLQ: eHealth Literacy Questionnaire
FEV1: forced expiratory volume during the first second
GOLD: Global Initiative for Chronic Obstructive Lung Disease

XSL•FO
RenderX

heiQ: Health Education Impact Questionnaire
HLQ: Health Literacy Questionnaire
PROM: patient-reported outcome measure
SF-36: Short Form-36

Edited by G Eysenbach; submitted 02.05.18; peer-reviewed by J Edwards, G Bravo; comments to author 04.10.18; revised version received 27.02.19; accepted 04.04.19; published 02.06.19.

Please cite as:
Lindskrog S, Christensen KB, Osborne RH, Vingtoft S, Phanareth K, Kayser L
Relationship Between Patient-Reported Outcome Measures and the Severity of Chronic Obstructive Pulmonary Disease in the Context of an Innovative Digitally Supported 24-Hour Service: Longitudinal Study
J Med Internet Res 2019;21(6):e10924
URL: https://www.jmir.org/2019/6/e10924/
doi:10.2196/10924
PMID:31199320

©Signe Lindskrog, Karl Bang Christensen, Richard H Osborne, Søren Vingtoft, Klaus Phanareth, Lars Kayser. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 02.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Creating Consumer-Generated Health Data: Interviews and a Pilot Trial Exploring How and Why Patients Engage

Kara Burns[1], BApp (Sci), BSci (Hons), PhD; Craig A McBride[2], BHB, MBChB, FRACS; Bhaveshkumar Patel[2], MBChB, FRACS; Gerard FitzGerald[3], MD, PhD, FACEM, FRACMA, FCHSM; Shane Mathews[1], BBus (Hons), PhD; Judy Drennan[1], BEd, MEd, PhD

[1]QUT Business School, Queensland University of Technology, Brisbane, Australia

[2]Queensland Children's Hospital, Children's Health Queensland, Brisbane, Australia

[3]School of Public Health & Social Work, Queensland University of Technology, Brisbane, Australia

Corresponding Author:
Kara Burns, BApp (Sci), BSci (Hons), PhD
QUT Business School
Queensland University of Technology
George St
Brisbane, 4000
Australia
Phone: 61 414294967
Email: drkaraburns@gmail.com

Abstract

Background: Consumer-generated health data (CGHD) are any clinically relevant data collected by patients or their carers (consumers) that may improve health care outcomes. Like patient experience measures, these data reflect the consumer perspective and is part of a patient-centric agenda. The use of CGHD is believed to enhance diagnosis, patient engagement, and thus foster an improved therapeutic partnership with health care providers.

Objective: The aim of this study was to further identify how these data were used by consumers and how it influences engagement via a validated framework. In addition, carer data has not been explored for the purpose of engagement.

Methods: Study 1 used interviews with CGHD-experienced patients, carers, and doctors to understand attitudes about data collection and use, developing an ontological framework. Study 2 was a pilot trial with carers (parents) of children undergoing laparoscopic appendectomy. For 10 days carers generated and emailed surgical site photographs to a tertiary children's hospital. Subsequently, carers were interviewed about the engagement framework. In total, 60 interviews were analyzed using theme and content analysis.

Results: This study validates a framework anchored in engagement literature, which categorizes CGHD engagement outcomes into 4 domains: physiological, cognitive, emotional, and behavioral. CGHD use is complex, interconnected, and can be organized into 10 themes within these 4 domains.

Conclusions: CGHD can instigate an ecosystem of engagement and provide clinicians with an enhanced therapeutic relationship through an extended view into the patient's world. In addition to clinical diagnosis and efficient use of health care resources, data offer another tool to manage consumers service experience, especially the emotions associated with the health care journey. Collection and use of data increases consumers sense of reassurance, improves communication with providers, and promotes greater personal responsibility, indicating an empowering consumer process. Finally, it can also improve confidence and satisfaction in the service.

(J Med Internet Res 2019;21(6):e12367) doi:10.2196/12367

KEYWORDS

patient generated health data; patient engagement; patient participation; mHealth; photography

XSL•FO
RenderX

Introduction

Consumer-generated health data (CGHD) is any clinically relevant data, collected and controlled by a patient or carer, to be used in the health care setting [1]. It has been dubbed the "blockbuster drug of the century" [2] because of the speculation that CGHD will promote engagement [3]. Research indicates CGHD improves quality of life [4], promotes health behaviors [5], increases considerations of service value and satisfaction [6,7], and has the potential to increase patient empowerment [8]. CGHD can be distinguished from patient-reported outcome measures, which are typically managed by providers [9], although both are now recognized to promote patient engagement and enhance care through monitoring and assessing symptoms, informing treatment decisions, tracking outcomes, agenda setting, and enhancing patient-provider communication [10].

CGHD is primarily useful to health care professionals in the diagnosis and management of conditions. Data can be clinician requested (solicited) or patient initiated (unsolicited) and are aggregated in many forms [1]. Quantitative measures of spirometry, qualitative descriptions such as a diary of emotions, and visual information such as photographs of dietary intake are common [11], with data capture encouraged through mobile phone apps and wearable technology. This research focuses on consumer-generated photographs that are useful for rheumatology [12], dermatology [4], surgical documentation [13], and wound management [14].

Research on CGHD has predominantly focused on benefits for the health care professional. The use and engagement outcomes promoted by the data for the patient and carer is still under-researched [15]. Previous studies show that data improves patient activation [16], a state characterized by an informed, knowledgeable, active patient who can sustain a course of health care under pressure [17]. Tang et al (2006) [16] suggest patients could experience a better quality of life and Frühauf et al (2012) [4] report this in a study of tele-dermatology where "all patients perceived they had made savings of time and expenses, and moreover, they believed they had gained a more flexible and empowered lifestyle." Tan et al (2012) conducted a qualitative interview study based on perceptions of unsolicited photographs. Notably, general practitioners also believed the images could empower patients, saying it helped patients retain control and "the patient became more part of the team [8]." In addition, carers are critical to engagement in some patient groups [18]; thus, their perspective is also valuable to explore.

A recent systematic review of patient engagement found that of the 89 randomized controlled trials that purported to instigate and measure engagement, 21 of those had no quantifiable measure. Of the 10 high-quality papers elected for review, only 1 study defined engagement a-priori [19]. Little research has evaluated the effect of CGHD on engagement, and an ontological framework for describing and measuring patient engagement is absent in extant literature [20]. Given the importance of patient engagement, the growing evidence of its ability to improve health outcomes, defining engagement is critical.

Using the education engagement literature, this study defines engagement as composed of physiological (clinical), cognitive, emotional, and behavior dimensions [21]. Engagement was chosen because it is a multifaceted, situation-specific concept, which can include a process and an end state. Focusing on consumers' individual experiences of CGHD and engagement as an end state, "outcome" is consistent with the patient centric aspirations of this study. Hence, this research explores an ontological engagement framework of cognitive, emotional, behavioral, and physiological outcomes for the use of consumer data to answer the research questions: *How are consumer generated health data used by consumers? How does this data influence engagement?*

Methods

Study 1 was conducted to explore experiences of CGHD, whereas study 2 was used to validate these findings in clinical care and explore carer-generated data. This 2-study research agenda improves construct validity through cumulative validation and confirmation by key informants [22,23] and clinician interviews used to validate consumer perspectives.

Study 1 sought a purposive sample of patients, carers, and doctors who were experienced in consumer-generated photography and used semistructured interviews to explore data use. Maximum variation sampling was applied by sampling from across the medical subspecialities of general practice, emergency health care, surgical care, and trainee. In addition, all consumers were judged as once-off users, intermittent users, regular users, and constant users of CGHD. This offered an opportunity to explore the widest possible attitudes, perceptions, and beliefs about the data via a cohort who understands the implications and complexity of its use. The sample size used in qualitative research relies on idea saturation to ensure reliability, and 30 participants was deemed suitable.

Study 2 utilized a pilot clinical trial with theoretical sampling of 30 parents of children undergoing laparoscopic appendectomy surgery at a large tertiary children's hospital. Using a standard operating procedure, parents were trained to take photographs of their children's surgical site wounds every 2 days and email them to the staff for the 10 days after the procedure. An all-comers approach consistent with other laparoscopic interventions was used [24]. Postintervention interviews were conducted by the researcher using introductory open and probing questions to deepen the understanding of CGHD use, engagement, and to validate the ontological framework.

Using grounded theory, interviews were analyzed using constant comparison for meaning using a 5-step process: (1) Gaining familiarity through reading all transcripts, (2) Data reduction via coding, (3) Thematic analysis of the codes, (4) Data reorganization into the 4 domains and comparison across both studies, and (5) A discussion of the key considerations related to the research question [25].

The 2 studies were analyzed sequentially but iteratively compared. Codes and themes from study 1 (perceptions of the CGHD) emerged from the data, then were then mapped to the ontological framework. Study 2 (use of data in care) provided

XSL•FO
RenderX

validation of the existing codes and themes and contributed new information. When new codes and themes emerged from study 2, they were compared and validated by study 1 findings and reorganized into the final framework. The combination of the 2 approaches ensured consumers' perceptions of CGHD use matched actual use in clinical care.

Results

Overview

Saturation was achieved at 60 interviews. In study 1, 19 males and 16 females participated in 34 interviews (Multimedia Appendix 1). Participants were aged between 22 to 69 years. A total of 16 patients and 7 carers were from a wide range of socioeconomic situations, with doctors from the 4 predefined categories. In study 2, 5 male and 21 female carers completed the trial by taking between 1 and 6 photographs of their child's surgical site wounds over 10 days (Multimedia Appendix 2). Of the original 30, 4 participants did not complete the trial. Participants were aged between 18 to 70 years, with children aged between 1 to 17 years.

Although there was significant similarity, study 1 and study 2 contributed unique findings to the results. Study 1 provided the 8 themes of improved health outcomes, self-perception, emotional regulation, empowerment preventative mind-set, self-management, social support, and partnership with providers, with study 2 contributing service optimization and service assessment. Although there were no major differences between patients and carers, doctors contributed the finding of patient deviance, which was supported then by carers.

In this context, CGHD has 30 main use outcomes grouped into 10 themes fitting within the ontological framework. Although outcomes have been placed into the physiological, cognitive, emotional, and behavioral groups, engagement is contextual and dynamically present in individual consumers. Outcomes were thus found to overlap and be interrelated, consistent with previous seminal research [21]. When the health care consumer perspective is analyzed, the physiological, cognitive, emotional, and behavioral outcomes of engagement are present but with a strong emphasis on the emotional outcomes of CGHD-related engagement.

Improved Health Outcomes

> *If you can take photos over time of visual things like rashes and growths and whatever, you can see whether they're getting worse or better. It's easy to miss subtle changes over time.* [Female Patient]

> *I also do think that there is a certainly much bigger potential in terms of research, and you know, into ongoing treatment or into the monitoring all sorts of illnesses.* [Male Carer]

Data had multiple, overlapping uses for physiological engagement including diagnosis, management, and medical research. Diagnosis and management of transient conditions, chronic conditions, slow healing, and slow-progressing conditions occurred with CGHD. This was enabled through the recording of changes and treatment over time. Doctors' opinions of the data were mixed, with it considered redundant information, partial information, and key information. A specific area of both opportunity and concern was remote diagnosis. Patients and carers agreed that remote diagnosis was possible with these data, indicating it overcame barriers of distance improving their quality of life, saving them time and money [4], although clinicians were skeptical about diagnosing without seeing the patient in person. In addition, CGHD can be used for research purposes, with the majority of discussion about projects led by health care staff. Patients were comfortable sharing CGHD, providing adequate consent is obtained. The caveat to this was that patient age, sensitivity of the condition, sensitivity of the body part, and anonymity were all factors in the consent process.

Self-Perception

> *With the breast cancer, the motivation [to take images] was this weird thing. I had to remember what I looked like before the surgery and come to terms with the fact that it's really happening to me. It's like a concrete object that reflects what's happening.* [Female Patient]

> *I think it [photography] makes me feel like a more responsible by engaging in that process...It definitely made me feel like you've been more responsible about the healing of the wound, without a doubt.* [Female Carer]

One of the most discussed themes in all interviews was the notion the data were evidence of a patient's personal record and of the medical experience. Recording the health care journey is a preliminary step toward health care sense-making and a reaction to the stressors of disease, which is intended to generate a sense of coherence [26]. In short, generating data is an expression of a patient's desire for control in the face of a stressful illness and helped consumers make sense of what was occurring. Patients documented their bodies before and after medical interventions as an act of both vanity and the documentation of survivorship [27]. Carers suggested the data were going to kept as mementos for adolescent children so that when they grow up they have a record about what they have gone through in their own lifetime.

Responsibility was a category that emerged as important for both patients and carers when discussing CGHD. Taking on extra responsibility through documentation was considered a consequence of data use. Increasing a patient's responsibility was suspected to improve health outcomes; however, doctors were wary with many concerned about this practice. One commented:

> *The disadvantage is that – I guess, the patient is now responsible for looking after the moles themselves, they may slightly be falsely assured in some way.* [Emergency Medicine Consultant]

Service Assessment

> *The opportunity to send those pictures and assess the risk of infection and things like that to a trained professional have a look like rather than making your own judgment or Googling. I think that's really*

XSL•FO
RenderX

> *important, and that's what I mean when I say satisfaction.* [Male Patient]
>
> *If that [remote photography assessment] became a regular thing I would have confidence in the system rather than the individual doctors. I would have more confidence that the problem has a chance of being resolved...none of the doctors are infallible.* [Male Carer]

Trust for the doctor, satisfaction with the service, and service confidence were all considered cognitive outcomes of the use of CGHD in clinical care; however, consumers did acknowledge trust was influenced by more than just using the data. Indeed, patients thought using the data may signal an increase in their trustworthiness and equate to more respect from the clinician in the service interaction. Doctors agreed that using data increased a sense of trust:

> *Patients feel that we trust them...it sort of creates a relationship where the patient is sure the doctor trusts this picture.* [Surgical Registrar]

In general, service assessment relies on evaluating the frontline service interactions through a prism of relational factors between the consumer and provider [28]. Satisfaction was present and discussed in interviews but remains an indistinct concept, conflated with service confidence, reassurance, and going beyond what is expected. Study 1 suggested CGHD could improve satisfaction, and in study 2, many parents commented that was indeed the case. One female carer was enthusiastic about her service experience, commenting that satisfaction was about "[when] those photos had gone back to the surgeon you know that that feels like going above and beyond the normal standard service."

Preventative Mind-Set

> *I was probably just made more aware of just healthcare in general by being involved in taking the photos.* [Female Carer]
>
> *They [wound photographs] just reminds me that I've got to watch what I'm doing. As the doctor explained to me when I was leaving the hospital, he says because I've got diabetic feet which is neuropathy in the feet, I've gotta' wear shoes all the time.* [Male Patient]

CGHD increased awareness of a condition and promoted health behaviors. This was most clearly expressed by 1 patient who noted photographs reminded her to reduce her stress levels to improve facial acne. Although it is clear that data can help awareness and serve as a reminder for healthy behaviors, recent literature suggests that although engagement using devices has the potential to facilitate health behavior change, this might not be driven by devices alone [29]. Commensurate with these findings, this research suggests data were a facilitator and prompt for behaviors, not a mechanism for change in itself.

In study 2, parents who documented their child's wounds used data collection as an excuse and reminder to check healing progress. It was discussed that adolescent children may not allow parents to observe wound healing, but the photography was "for the doctor" and therefore permitted. Interestingly, participants used the photography as a prompt to self-manage care episodes, with 1 parent saying:

> *It was really good for me because otherwise my management of time can be quite poor sometimes and I would - may forget to - I probably would have forgotten to check his wound.* [Female Parent]

Emotional Regulation

> *I did that [taking photographs] around the time I started caring for Dad. You start noticing all these other things when you start caring for someone, so it also worked well for my own peace of mind. It's reassurance, yes they will believe that yes this has occurred, and this is ongoing.* [Female Carer]
>
> *I was going to take a photo of my butt because I had a big bruise there, but that wasn't for any, it wasn't for any kind of diagnosis or anything. That was just posting on social media because, you know, getting a response from people.* [Male Patient]

CGHD provided reassurance by helping consumers cope with difficult emotions, and this phenomenon was recognized by all 3 participant groups. Patients and carers used data for increased emotional reassurance that healing was occurring and that they were performing the required tasks to get better. In opposition to reassurance, anxiety was suggested as creating data might cause undue stress. Interestingly, both doctors and patients recommended that capturing data may increase the anxiety related to the condition, and hence, not all participants would be suitable for data collection. Indeed, although data collection may be a sign of diligence in a patient, it might also be a sign of hypochondria, tracking conditions that are not required, which in turn cause distress.

CGHD offered consumers an avenue to share data with others as a form of entertainment within the peer network and to improve the lives of others through altruistic acts. Despite being characterized as "grotesque" and "embarrassing," male participants used the graphic nature of medical images for play and entertainment of others. Patients were also motivated to share data by a feeling of altruism, giving information to benefit the wider community, which is commensurate with Spencer et al [30] who reported 98% of participants who shared anonymous health information for research purposes "considered that the altruistic benefits of sharing health care data outweighed the risks."

Empowerment

> *[With videos] the doctors can see that I'm not making this thing up. I had a history of depression before so they would say, "well, funny things can happen when you get depressed." So, it was valuable for me to be able to show it to someone, like "here it is"… I became more of an advocate myself.* [Male Patient]
>
> *When the patient takes the photograph, the patient themselves has thought, "Oh this is something. I need to do something about it and actually pick up the camera." To do that gives you a little bit of sense of control.* [Female Carer]

XSL•FO
RenderX

The most important finding of emotional engagement is that generating data was an empowering process promoting self-advocacy, self-confidence, and a feeling of control. A patient with a history of depression suffered intermittent hand twitching generated medical videos for a diagnosis. As such, the data provided an avenue for self-advocacy, giving the consumer more credibility with health care professionals. Self-advocacy was also present when carers used data to assert their role within families. Data focused conversations with family, making the process of healing easier and reassuring the carers that patients were improving. Data also promoted self-confidence, suggesting they were functioning as a "good" carer.

Furthermore, establishing a sense of control was important to many consumers, who felt health care services can be "dehumanizing" and "disempowering." Generally, the sense of control was expressed as both health situation and health system control. CGHD use improved the patient's feeling of equality with the health care provider and helped access health care services in a timely manner equating to a sense of control over the health care service. At home, it reassured patients and carers that they had a role in managing care, improving their perception of control over the health situation.

Self-Management

> *The main thing I wanted to track was the rate at which it was healing. So, I used to take photographs, along with a ruler on the side of the wound and that way, we could actually pick up, at some point that the wound had arrested.* [Male Patient]

> *If you wanted to do some self-diagnosis, you've got the proof sitting there behind you of these photos.* [Female Parent]

Even before digital devices, patients were offered opportunities to monitor their own condition. This occurred for chronic diseases such as diabetes and improved self-efficacy, even in patients from diverse backgrounds [31]. It was universally observed by patients, carers, and doctors that digitally created CGHD is used for self-monitoring and self-education, which can lead to the self-diagnosis. Patients considered the use of Dr Google a health care right but realized that not all information is accessible or reliable. Finding a legitimate source for self-education was considered difficult because of multiple illegitimate sources, and the readability of most authoritative Web pages exceeds average national levels of literacy [32].

Doctors stressed that although the data may be useful for self-diagnosis, they can lead to misdiagnosis and misinformation. Misinformation has medico-legal implications for health care professionals, and although 1 doctor commented they would never be "stupid enough" to diagnose from consumer-generated data, others acquiesced, suggesting that treating the patient also means treating data and the ideas they present with. Refusal to engage with patients who present data will impact the service experience. Indeed, patients who present data-evidenced opinions switched doctors if their perspectives were not recognized.

Social Support

> *I think it's a bit of a connection thing. If people do see it [a photograph] and they say "Oh," you know, "I can sympathise," and I think that's what the medical profession is lacking.* [Male Patient]

> *I have a colleague at work who is also a very good friend. She's having surgery on Monday I was glad I had my daughter's photos and was able to send my friend at work the progressive photos to show her how well the wounds were healing.* [Male Carer]

Data has an important function in the social support of patients with both data sharing and information seeking, improving a sense of connectedness through internet surrogates. Data sharing in online and offline networks for social support was very common. For example, an older female patient was supported through her knee reconstruction and encouraged to continue treatment with supportive comments that she was "doing well" and to "keep up the good work." This is concordant with literature that suggests social networks support adherence to short duration activity regimes [33].

CGHD was shared with the patient network for an update on the patient status and information seeking in study 1. Notably, no information seeking was observed in study 2, disconfirming the use of the data in this context. When participants were asked why, approximately half mentioned the absence of this behavior, suggesting they did not need to search online or use the peer network for information seeking, potentially reducing instances of self-diagnosis. The reasons given were both that a diagnosis was already made, and that because of the remote support of the doctors', self-diagnosis was not required.

Partnership With Providers

> *By the time I saw my doctor I had a skin graft taken and applied to my arm, so it looked pretty normal. However, the photos after the operation, when I was changing the dressing, pretty much showed an arm that was cut from elbows to palm with all the muscles sort of hanging out, sitting on the table. And that it would have been impossible for the doctor to understand or see that scenario without taking photos.* [Male Patient]

> *I took a photo and in a series of about 3 hours, you could see the breakdown of it [the wound] happening; we were able to use it as a thing, to say this needs urgent attention now…I bumped into one of the doctors and she told the surgeon…I think I showed her the photo of it breaking down and she called the surgeon, and the surgeon stepped in.* [Female Carer]

CGHD instigated cooperation between consumers and providers, improved communication, and gave the provider an opportunity to offer parents support by distance and equated to more perceived respect for the consumer. Improved patient–provider communication was a well-documented outcome of data use [34]. It was universally discussed by patients, carers, and doctors that a feature of photographic data was that it overcame an inability to describe a condition. Furthermore, when the data were used in clinical care, it could be used as a focus of

XSL•FO
RenderX

conversation or for agenda setting. It prompted questions and feedback with participants commenting the data improved meeting participation with health care providers. Importantly, the communication was bidirectional, with clinicians also experiencing the benefit of improved communication, suggesting it prompted other questions they might not have asked.

In addition to improved communication, cooperation was noted in interviews. Data were used to educate and motivate clinicians, increasing a sense of urgency for the treatment of patient conditions. Images were not only used as a "call to action" between the patient and doctor, but clinicians used patient data to convince their colleagues to urgently treat their patients. Consumers understood the "power relationship in a medical model," however, they sought cooperation with clinicians when using CGHD. It was suggested that health care professionals could cooperate with patients by training them on the data capture, and clinicians agreed it was part of their job to educate patients. If cooperation was not enacted, for example, by "shrugging off" consumer data, it caused dissatisfaction with the service experience.

Consumer-generated data affected parents' perceptions of the service through supported autonomy at home. Patients and carers experienced a sense of autonomy in their health situation, however, are still supported by health care professionals within the health system through remote diagnosis of the images. This helped rebuke the feeling of "having the door swing shut on you once you have left the service" and was considered going above and beyond what was expected, extending the service relationship beyond the clinical context.

Finally, a regular theme prominent in study 2 was parents recognizing that through the act of providing data, they gained greater attention and more respect. Typically, this was expressed as being taken more seriously. Participants attributed this to being more aware of the health condition, proactive when caring, and taking on greater responsibility for their child's health. A total of 1 patient experienced greater respect when she alerted medical staff to an infection discovered through consumer-generated photography:

> *I think that they probably may take you a little bit more seriously because you said you were aware that you know 12 hours ago that wasn't there, or six hours ago that wasn't there, and then suddenly it was something like we have maybe a bit more respect for parents.* [Female Parent]

Service Optimization

> *You can monitor things yourself. If you notice a significant change in 6 months' time, because you're looking at yourself and you can compare it with the original photo that was taken. So, you might go back to the doctor in 6 months, rather than waiting the 12 months for your scheduled check-up.* [Male Patient]

> *[Using photographs] people might have demanded to see a surgeon straight away, or, whatever and I probably just tried to stay calm, and just say, you know, "when you got to me that you just popped in here and have a look at this."* [Female Carer]

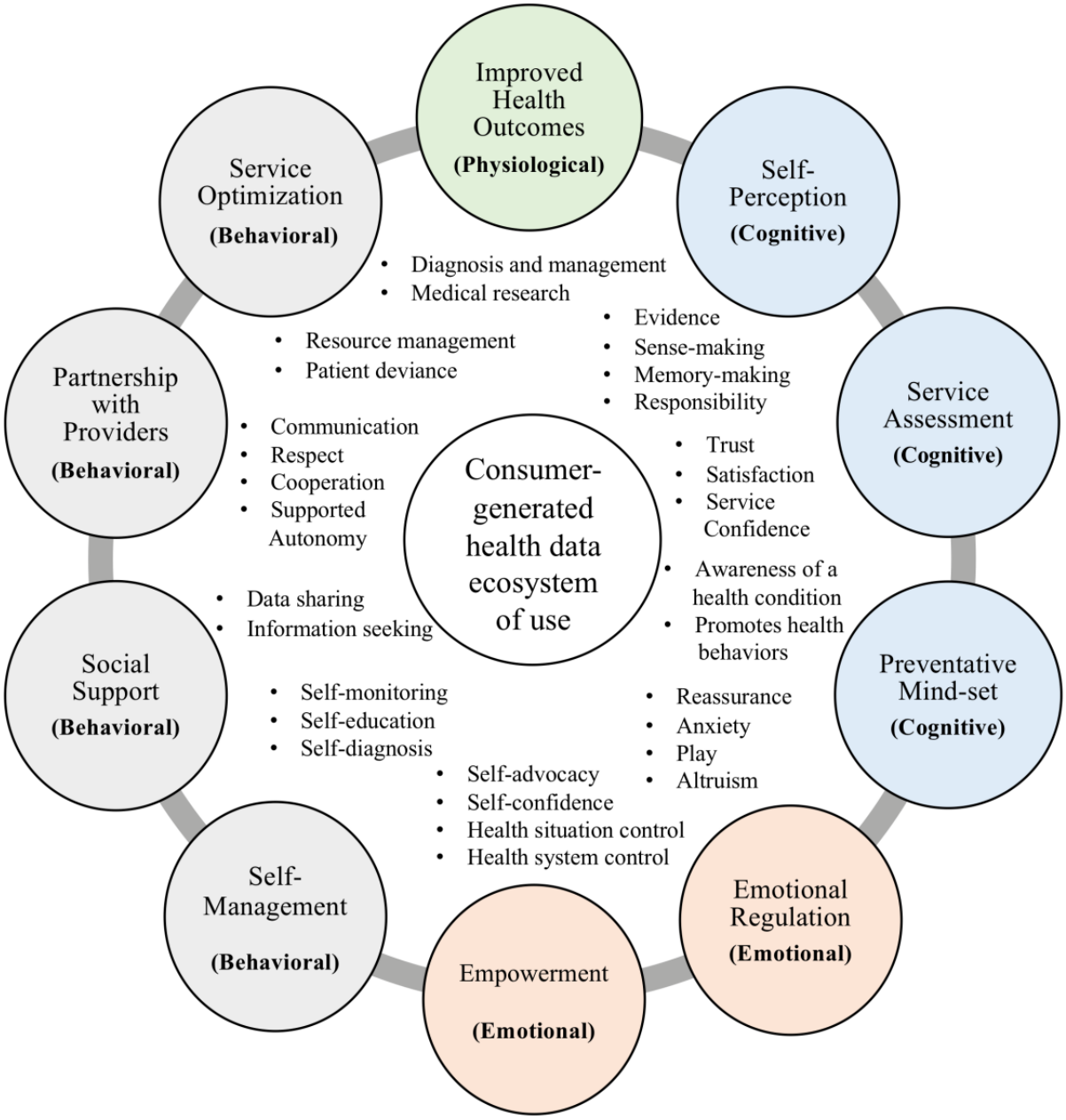
Patients and carers experienced a greater ability to effectively manage their own resources and reduce health care services use by eliminating unnecessary appointments. The main way patients used data in this context was to alter their own treatment plan, but this could also lead to deviant behaviors. Typically, this meant reviewing photographs for evidence of clinical improvement and then making appointment changes based on symptoms shown in the data. Consumers commented the data could provide detection of adverse health events, motivating an early visit to a health care service. Doctors, patients, and carers all agreed that getting timely appointments for transient conditions was difficult and that health services lacked continuity between appointments, with CGHD overcoming these limitations. Finally, when clinician's expertise was limited, CGHD could be sent to experts for review.

Deviant customer behavior refers to actions that patients take to abuse the health care system that violate accepted norms of behavior and result in harm [35]. The research found no evidence that patients will take actions that affect other patients but may exhibit deviant behaviors toward the health care provider. Carers and doctors, but not patients, agreed that deviant behavior involved asking for diagnosis based on the data for someone not present; self-diagnosis, which delays seeking medical help; and insisting that the patient's interpretation of the data is correct despite contradicting the clinical expert. A total of 4 doctors commented that when patients use the data for self-diagnosis, they often come to the consultation with a preconception about what they have. When diagnosis is inconsistent with expectations, patients may get a second opinion and change providers.

As clinical costs rise, private telehealth offerings emerge, and health tracking gains popularity. Providers face unprecedented pressure to develop cheaper, patient-centric, value-based health services. Engaging patients has become a key focus patient-centered care and has the potential to improve both the quality and efficiency of health care services. CGHD can improve service accessibility and promote engagement. Thus, Figure 1 demonstrates CGHD use attributes go beyond the physiological gains of improved health outcomes to incorporate an ecosystem of physiological, cognitive, emotional, and behavioral engagements (also see Multimedia Appendix 3).

XSL•FO
RenderX

Figure 1. An engagement framework for consumer-generated health data.



Discussion

Principal Findings

> *Having those medical photos going to a medical practitioner has meant that I felt more and more engaged, more involved and more responsible. I was given some sort of capability to assist in the process.* [Female Carer]

Consumers use data for improved health outcomes, self-perception, service assessment, emotional regulation, empowerment, social support, partnership with providers, and service optimization. For providers, data can aid diagnosis and management, improve communication, and can reduce unnecessary consultations. CGHD is another tool to manage the patient journey, allowing providers an opportunity to assure consumers that healing is occurring. For consumers, data aides sense making, instigates greater personal responsibility for health care outcomes, can foster improved awareness of a condition, and promotes healthy behaviors.

Limitations

The concepts demonstrated have been explored qualitatively, and no quantitative confirmation of relationships was undertaken. In addition, many of the themes and uses overlap, and some may be categorized into more than 1 area. The subjective nature of qualitative research does introduce the possibility of researcher bias; however, the techniques of intercoder reliability, maximum variation sampling, and a theoretical framework developed from the patient engagement literature were used to mitigate these issues. A further limitation is that patients may have very different perspectives on CGHD use, with other technologies and clinical scenarios thus impacting on engagement.

XSL•FO
RenderX

Comparison With Prior Work

Although patients and carers wholeheartedly supported data use, 1 clinician refused to see the data at all. This confirms other findings that a small percentage of health care practitioners will not be interested CGHD and may miss the opportunity to improve the therapeutic relationship [3,36]. Importantly, this study confirms and explains why empowerment occurs, suggesting it is an experience of self-advocacy, self-confidence, health system control, and health situation control and more than just "being part of the team" [8].

Conclusions

Today, consumers are connecting in online and offline networks, collating health information through mobile technology, and looking for innovative services that improve their participation. Thus, managers should aim to develop strategies to promote CGHD, as refusal to engage with patients who present data will impact service experience. Indeed, our research shows patients who present data-evidenced opinions will switch doctors if their autonomy is not recognized.

Furthermore, patients are much more likely to attempt to self-diagnose using consumer-generated data when the doctor was not available and when they have a peer network. Therefore, "prescribing" data collection may reduce instances of self-diagnosis. If data are supported then it can lead to higher service satisfaction, higher service confidence, and an improved therapeutic relationship. Finally, future research should address the use of other data in other clinical settings adding to the existing theoretical framework limited to acute and chronic health care.

Authors' Contributions

All authors contributed to the design on the research. KB was the principal investigator and author. All authors contributed to the review of the final draft of the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Study 1 participant characteristics.

[DOCX File, 29KB - jmir_v21i6e12367_app1.docx]

Multimedia Appendix 2

Study 2 participant characteristics.

[DOCX File, 17KB - jmir_v21i6e12367_app2.docx]

Multimedia Appendix 3

Themes and participant quotes.

[DOCX File, 23KB - jmir_v21i6e12367_app3.docx]

References

1. Shapiro M, Johnston D, Wald J, Mon D. HealthIT. 2012 Apr. Patient-Generated Health Data URL: https://www.healthit.gov/sites/default/files/rti_pghd_whitepaper_april_2012.pdf [accessed 2019-05-26] [WebCite Cache ID 78esc4NDN]
2. Taylor L. Preparing for patient-generated documents. J AHIMA 2013;84(5):46. [Medline: 23755434]
3. Ahern DK, Woods SS, Lightowler MC, Finley SW, Houston TK. Promise of and potential for patient-facing technologies to enable meaningful use. Am J Prev Med 2011 May;40(5 Suppl 2):S162-S172. [doi: 10.1016/j.amepre.2011.01.005] [Medline: 21521591]
4. Frühauf J, Schwantzer G, Ambros-Rudolph CM, Weger W, Ahlgrimm-Siess V, Salmhofer W, et al. Pilot study on the acceptance of mobile teledermatology for the home monitoring of high-need patients with psoriasis. Australas J Dermatol 2012 Feb;53(1):41-46. [doi: 10.1111/j.1440-0960.2011.00852.x] [Medline: 22309330]
5. McClellan S, Panattoni L, Chan A, Tai-Seale M. Patient-initiated electronic messages and quality of care for patients with diabetes and hypertension in a large fee-for-service medical group: results from a natural experiment. Med Care 2016 Mar;54(3):287-295. [doi: 10.1097/MLR.0000000000000483] [Medline: 26759978]
6. Klein D, Fix G, Hogan TP, Simon SR, Nazi K, Turvey CL. Use of the blue button online tool for sharing health information: qualitative interviews with patients and providers. J Med Internet Res 2015 Aug 18;17(8):e199 [FREE Full text] [doi: 10.2196/jmir.4595] [Medline: 26286139]
7. Baer D. Patient-physician e-mail communication: the kaiser permanente experience. J Oncol Pract 2011 Jul;7(4):230-233 [FREE Full text] [doi: 10.1200/JOP.2011.000323] [Medline: 22043186]

XSL•FO
RenderX

8. Tan L, Hu W, Brooker R. Patient-initiated camera phone images in general practice: a qualitative study of illustrated narratives. Br J Gen Pract 2014 May;64(622):e290-e294 [FREE Full text] [doi: 10.3399/bjgp14X679723] [Medline: 24771843]
9. Murthy HS, Wood WA. The value of patient reported outcomes and other patient-generated health data in clinical hematology. Curr Hematol Malig Rep 2015 Sep;10(3):213-224. [doi: 10.1007/s11899-015-0261-6] [Medline: 26040262]
10. Lavallee DC, Chenok KE, Love RM, Petersen C, Holve E, Segal CD, et al. Incorporating patient-reported outcomes into health care to engage patients and enhance care. Health Aff (Millwood) 2016 Apr;35(4):575-582. [doi: 10.1377/hlthaff.2015.1362] [Medline: 27044954]
11. Casperson SL, Sieling J, Moon J, Johnson L, Roemmich JN, Whigham L. A mobile phone food record app to digitally capture dietary intake for adolescents in a free-living environment: usability study. JMIR Mhealth Uhealth 2015 Mar;3(1):e30 [FREE Full text] [doi: 10.2196/mhealth.3324] [Medline: 25775506]
12. Armstrong D. The mobile phone as an imaging tool in SLE. Rheumatology (Oxford) 2004 Sep;43(9):1195. [doi: 10.1093/rheumatology/keh286] [Medline: 15317959]
13. Gunter R, Fernandes-Taylor S, Mahnke A, Awoyinka L, Schroeder C, Wiseman J, et al. Evaluating patient usability of an image-based mobile health platform for postoperative wound monitoring. JMIR Mhealth Uhealth 2016 Sep 28;4(3):e113 [FREE Full text] [doi: 10.2196/mhealth.6023] [Medline: 27683059]
14. Michiels E, Ouellette L, Bush C, VanDePol E, Fleeger T, Jones JS. Camera phones for the follow-up of soft-tissue injuries in adult and pediatric ED patients: a feasibility study. Am J Emerg Med 2017 Dec;35(5):782-783. [doi: 10.1016/j.ajem.2016.10.033] [Medline: 28188056]
15. Sands D, Wald J. Transforming health care delivery through consumer engagement, health data transparency, and patient-generated health information. Yearb Med Inform 2014 Aug 15;9:170-176 [FREE Full text] [doi: 10.15265/IY-2014-0017] [Medline: 25123739]
16. Tang PC, Ash JS, Bates DW, Overhage JM, Sands DZ. Personal health records: definitions, benefits, and strategies for overcoming barriers to adoption. J Am Med Inform Assoc 2006;13(2):121-126 [FREE Full text] [doi: 10.1197/jamia.M2025] [Medline: 16357345]
17. Hibbard J, Stockard J, Mahoney E, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
18. Tuffrey-Wijne I, Giatras N, Goulding L, Abraham E, Fenwick L, Edwards C, et al. National Institute for Health Research. 2013. Identifying the factors affecting the implementation of strategies to promote a safer environment for patients with learning disabilities in NHS hospitals: a mixed-methods study URL: https://www.journalslibrary.nihr.ac.uk/hsdr/hsdr01130/ [accessed 2019-06-01] [WebCite Cache ID 78olkH3GX]
19. Simmons LA, Wolever RQ, Bechard EM, Snyderman R. Patient engagement as a risk factor in personalized health care: a systematic review of the literature on chronic disease. Genome Med 2014;6(2):16 [FREE Full text] [doi: 10.1186/gm533] [Medline: 24571651]
20. Prey JE, Woollen J, Wilcox L, Sackeim AD, Hripcsak G, Bakken S, et al. Patient engagement in the inpatient setting: a systematic review. J Am Med Inform Assoc 2014;21(4):742-750 [FREE Full text] [doi: 10.1136/amiajnl-2013-002141] [Medline: 24272163]
21. Fredricks JA, Blumenfeld P, Paris AH. School engagement: potential of the concept, state of the evidence. Rev Educ Res 2016 Jul;74(1):59-109. [doi: 10.3102/00346543074001059]
22. Babbie E. The Basics of Social Research. Belmont: Wadsworth, Cengage Learning; 2013.
23. Sarantakos S. Social Research. Sydney: Macmillan International Higher Education; 2012.
24. Thompson MH, Tranter SE. All-comers policy for laparoscopic exploration of the common bile duct. Br J Surg 2002 Dec;89(12):1608-1612. [doi: 10.1046/j.1365-2168.2002.02298.x] [Medline: 12445074]
25. Tolley E, Ulin P, Mack N, Robinson E, Succop S. Qualitative Methods in Public Health: A Field Guide for Applied Research. New York: Wiley; 2016.
26. Antonovsky A. Unraveling the mystery of health: how people manage stress and stay well. In: The Health Psychology Reader. Thousand Oaks, CA: Sage Publications; 2002:127.
27. Chou W, Hunt Y, Folkers A, Augustson E. Cancer survivorship in the age of YouTube and social media: a narrative analysis. J Med Internet Res 2011 Jan 17;13(1):e7 [FREE Full text] [doi: 10.2196/jmir.1569] [Medline: 21247864]
28. Howden C, Pressey A. Customer value creation in professional service relationships: the case of credence goods. The Service Industries Journal 2008 Sep 16;28(6):789-812. [doi: 10.1080/02642060801990361]
29. Patel M, Asch D, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. J Am Med Assoc 2015 Feb 03;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
30. Spencer K, Sanders C, Whitley E, Lund D, Kaye J, Dixon W. Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study. J Med Internet Res 2016 Apr 15;18(4):e66 [FREE Full text] [doi: 10.2196/jmir.5011] [Medline: 27083521]
31. Sarkar U, Fisher L, Schillinger D. Is self-efficacy associated with diabetes self-management across race/ethnicity and health literacy? Diabetes Care 2006 Apr;29(4):823-829. [doi: 10.2337/diacare.29.04.06.dc05-1615] [Medline: 16567822]

XSL•FO
RenderX

32. Cheng C, Dunn M. Health literacy and the Internet: a study on the readability of Australian online health information. Aust N Z J Public Health 2015 Aug;39(4):309-314. [doi: 10.1111/1753-6405.12341] [Medline: 25716142]
33. Hunter R, McAneney H, Davis M, Tully M, Valente T, Kee F. "Hidden" social networks in behavior change interventions. Am J Public Health 2015 Mar;105(3):513-516. [doi: 10.2105/AJPH.2014.302399] [Medline: 25602895]
34. Huba N, Zhang Y. Designing patient-centered personal health records (PHRs): health care professionals' perspective on patient-generated data. J Med Syst 2012 Dec;36(6):3893-3905. [doi: 10.1007/s10916-012-9861-z] [Medline: 22644130]
35. Solomon M, Russell-Bennett R, Previte J. Consumer Behaviour : Buying, Having, Being. Frenchs Forest: Pearson Higher Education AU; 2012.
36. Nundy S, Lu CE, Hogan P, Mishra A, Peek ME. Using patient-generated health data from mobile technologies for diabetes self-management support: provider perspectives from an Academic medical center. J Diabetes Sci Technol 2014 Jan;8(1):74-82 [FREE Full text] [doi: 10.1177/1932296813511727] [Medline: 24876541]

Abbreviations

CGHD: consumer-generated health data

Edited by G Eysenbach; submitted 29.09.18; peer-reviewed by A Chan, E Austin; comments to author 15.11.18; revised version received 24.01.19; accepted 24.03.19; published 13.06.19.

Please cite as:
Burns K, McBride CA, Patel B, FitzGerald G, Mathews S, Drennan J
Creating Consumer-Generated Health Data: Interviews and a Pilot Trial Exploring How and Why Patients Engage
J Med Internet Res 2019;21(6):e12367
URL: http://www.jmir.org/2019/6/e12367/
doi:10.2196/12367
PMID:31199312

©Kara Burns, Craig A McBride, Bhaveshkumar Patel, Gerard FitzGerald, Shane Mathews, Judy Drennan. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based Mental Health Platform for Individuals Seeking Specialized Mental Health Care Services: Multicenter Pragmatic Randomized Controlled Trial

Jennifer M Hensel[1,2,3,4], MSc, MD; James Shaw[1,4], PhD; Noah M Ivers[1,4,5,6], MD, PhD; Laura Desveaux[1,4,6], PhD; Simone N Vigod[1,2,4,6], MSc, MD; Ashley Cohen[7], MSc; Nike Onabajo[1], MSc; Payal Agarwal[1,5], MD; Geetha Mukerji[1,6,8], MSc, MD; Rebecca Yang[1], MPH; Megan Nguyen[1], MN; Zachary Bouck[1], MSc; Ivy Wong[1], MPA, MPAff; Lianne Jeffs[7], RN, PhD; Trevor Jamieson[1,8], MD, MBI; R Sacha Bhatia[1,4,8], MD, MBA

[1]Women's College Institute for Health Systems Solutions and Virtual Care, Toronto, ON, Canada

[2]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

[3]Department of Psychiatry, University of Manitoba, Winnipeg, MB, Canada

[4]Women's College Research Institute, Toronto, ON, Canada

[5]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

[6]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[7]Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, ON, Canada

[8]Department of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jennifer M Hensel, MSc, MD
Women's College Institute for Health Systems Solutions and Virtual Care
76 Grenville St
Toronto, ON, M5S 1B2
Canada
Phone: 1 416 323 6400 ext 5126
Fax: 1 416 323 6004
Email: jennifer.hensel@wchospital.ca

Abstract

Background: Web-based self-directed mental health applications are rapidly emerging to address health service gaps and unmet needs for information and support.

Objective: The aim of this study was to determine if a multicomponent, moderated Web-based mental health application could benefit individuals with mental health symptoms severe enough to warrant specialized mental health care.

Methods: A multicenter, pragmatic randomized controlled trial was conducted across several outpatient mental health programs affiliated with 3 hospital programs in Ontario, Canada. Individuals referred to or receiving treatment, aged 16 years or older, with access to the internet and an email address, and having the ability to navigate a Web-based mental health application were eligible. A total of 812 participants were randomized 2:1 to receive immediate (immediate treatment group, ITG) or delayed (delayed treatment group, DTG) access for 3 months to the Big White Wall (BWW), a multicomponent Web-based mental health intervention based in the United Kingdom and New Zealand. The primary outcome was the total score on the Recovery Assessment Scale, revised (RAS-r) which measures mental health recovery. Secondary outcomes were total scores on the Patient Health Questionnaire-9 item (PHQ-9), the Generalized Anxiety Disorder Questionnaire-7 item (GAD-7), the EuroQOL 5-dimension quality of life questionnaire (EQ-5D-5L), and the Community Integration Questionnaire. An exploratory analysis examined the association between actual BWW use (categorized into quartiles) and outcomes among study completers.

Results: Intervention participants achieved small, statistically significant increases in adjusted RAS-r score (4.97 points, 95% CI 2.90 to 7.05), and decreases in PHQ-9 score (−1.83 points, 95% CI −2.85 to −0.82) and GAD-7 score (−1.55 points, 95% CI −2.42 to −0.70). Follow-up was achieved for 55% (446/812) at 3 months, 48% (260/542) of ITG participants and 69% (186/270) of DTG participants. Only 58% (312/542) of ITG participants logged on more than once. Some higher BWW user groups had significantly greater improvements in PHQ-9 and GAD-7 relative to the lowest use group.

XSL•FO
RenderX

Conclusions: The Web-based application may be beneficial; however, many participants did not engage in an ongoing way. This has implications for patient selection and engagement as well as delivery and funding structures for similar Web-based interventions.

Trial Registration: ClinicalTrials.gov NCT02896894; https://clinicaltrials.gov/ct2/show/NCT02896894 (Archived by WebCite at http://www.webcitation.org/78LIpnuRO)

(J Med Internet Res 2019;21(6):e10838) doi:10.2196/10838

KEYWORDS

internet; mental health; anxiety; depression

Introduction

Mental illness is prevalent, with estimates suggesting that upward of 1 billion people worldwide could be affected at a given point in time [1]. In addition, mental and substance use disorders are emerging as a leading cause of disability, accounting for nearly 10% of global disability-adjusted life years [1]. Access to and use of appropriate and timely mental health services and specialists, however, continues to be a challenge because of limited resources and individual-level factors surrounding treatment seeking [2-4]. E-mental health applications can potentially help to address some of these gaps [5]. A number of Web-based interventions including smartphone apps and Web-based treatment programs for common mental disorders have demonstrated small-to-moderate treatment effect sizes for symptom reduction [6-8], although these interventions are commonly recommended as standalone or preventative treatment options for those with milder symptoms, where benefits have been most apparent. Engagement with self-directed, Web-based interventions has been cited as a challenge owing to a range of user and intervention design factors, with multicomponent interventions potentially enhancing engagement through more user choice, added interactivity, and customization [9,10]. Moreover, the general advancement and adoption of virtual care is often in the absence of rigorous evaluation and adequate planning for sustainability and spread [11,12].

Investments in digital health worldwide have included a substantial emphasis on digital and virtual intervention to promote health [13], and Canada is no exception [14]. As a part of a series of demonstration projects being implemented by leading stakeholders in digital health and telemedicine in Ontario, Canada's most populous province, the Big White Wall (BWW) was selected as a solution with the potential to be adopted for mental health. The BWW [15] is a multicomponent moderated internet-based intervention with peer support that provides anonymity. At the time this trial was conducted, there had been no previous randomized trial evaluating the BWW.

The target goal for the Ontario demonstration projects was chronic disease management and high-needs patient populations, so the sponsors intentionally selected specialized mental health treatment settings for the intervention. This study sought to determine the utility of the BWW as a solution in a Canadian setting, specifically for individuals with mental health symptomatology severe enough to warrant a need for specialized mental health care. We investigated the effectiveness of 3 months of access to the BWW for mental health recovery, as well as symptoms of depression and anxiety, quality of life, and integration with one's community, relative to a usual care control group who received delayed access to the intervention after the study period. We hypothesized that the BWW would increase mental health recovery across a variety of mental health–related needs and conditions .

Methods

Trial Design

This study was a multicenter, parallel-arm, pragmatic randomized controlled trial. Participants seeking services at specialized mental health and addiction programs at the participating sites were randomized 2:1 to receive immediate access to the BWW for 3 months (immediate treatment group, ITG) or delayed access after a 3-month control period (delayed treatment group, DTG). The trial protocol has been previously published and is available open access [16]. The trial was sponsored by Ontario Telemedicine Network and Canada Health Infoway, both government-funded organizations. Sponsors specified the recruitment target and study settings, but had no involvement in the study design, procedures, data collection, or analysis.

Ethics

This study received ethical approval from the research ethics boards at all participating sites. All participants gave informed consent before taking part in the study.

Changes to Trial Design After Commencement

In response to automated follow-up surveys going to email *junk boxes*, personalized emails were sent with the survey link embedded. The frequency of contact to remind participants to complete follow-up surveys was increased, and surveys were completed over the phone when possible. To meet sponsor recruitment targets, recruitment was extended early in the study to individuals attending clinics rather than the initial approach targeting those on waitlists or being discharged. Given a large amount of referral between programs, early on in recruitment, it became apparent that participants were commonly on waitlists for specialized clinics while receiving active treatment in other recruitment settings, so this change had little impact on the overall composition of the study sample.

Recruitment of Participants and Baseline Assessment

Participants were recruited from outpatient mental health programs affiliated with 3 hospitals in Ontario: (1) a public psychiatric hospital in a medium-sized city with satellite

XSL•FO
RenderX

treatment sites in smaller urban centers; (2) a large community hospital located in another medium-sized city with satellite addictions programs; and (3) a large ambulatory academic hospital located centrally in a large metropolitan area. The aim of the study was to determine the utility of the BWW as a broad reaching multicomponent Web-based intervention that crosses all mental health and addiction-related needs. The BWW contains guided educational and course content applicable to a wide range of mental health needs, and the peer support component offers various opportunities to engage in discussion relevant to one's personal needs. Thus, we recruited from a range of outpatient programs which included: adult mood and anxiety psychiatry programs, a substance use program, an emergency department, an urgent care clinic, youth mood and anxiety programs, mood and anxiety psychotherapy programs, trauma therapy programs, and a borderline personality disorder program. Program coordinators, clinicians, and administrators at each site reviewed referral waitlists and clinic rosters to identify potential participants and refer them to the study coordinator, who reached out by telephone or met the individual at the clinic. Eligible participants were aged 16 years or older, had access to the internet and an email address, were able to read English, and willing and able to access and use a Web-based mental health intervention. There were no exclusion criteria but referring clinicians were asked to consider if person would be able to participate appropriately in Web-based peer interactions. There were no imposed restrictions on the use of concomitant care, including accessing other Web-based interventions that may be like the one under investigation.

All participants provided informed consent and baseline variables either in person or by phone with a study team member. Baseline assessment included sociodemographic data and all outcome measures. Participants were also asked about baseline belief in the treatment credibility and outcome expectancy. Participants were asked to rate their agreement with the author-generated question, "Self-help tools including web-based services and books are helpful for people with mental health problems." Item 4 from the Credibility and Expectancy Questionnaire [17] was adapted for this study, "By the end of the [BWW access period], how much improvement in your symptoms do you think will occur?" Participants were asked to rate their response from 0% to 100% with options available in 10% increments. This single item has been shown to correlate strongly with mental health treatment outcomes such as psychotherapy [18]. Participants were given the option to complete additional symptom, function, and service utilization measures over the phone or by computerized survey. Secure computerized assessments were sent electronically to each participant's email address. All data were entered into a REDCap database [19]. Trial recruitment took place between July 2016 and January 2017.

Participant Safety

The team at each study site monitored adverse events. A data safety and monitoring board consisting of 3 experts external to the research team reviewed the trial data at 3 and 6 months after recruitment started, specifically examining change in item 9 (suicidal ideation) scores on the Patient Health Questionnaire-9 item (PHQ-9), as well as all serious adverse events reported by the participating sites throughout the trial.

Intervention—The Big White Wall

The BWW is a multicomponent, self-directed, and moderated Web-based intervention founded in 2007 and hosted in the United Kingdom. Access to the BWW was free for study participants for 3 months (ITG during the 3-month trial period and DTG after the initial 3-month trial period). The developers advocate that the BWW can improve mental health symptoms through increasing social engagement, normalizing experiences, educating, and equipping with skills to manage difficulties. All participants maintain anonymity on the site through a unique nonidentifiable user alias and use is participant dependent [15]. The BWW is monitored 24/7 by *Wall Guides* who are trained mental health professionals employed by the BWW and based in the United Kingdom and New Zealand. These individuals constantly review user activity and posts to ensure the content is appropriate and sensitive to all users. They will engage with users through instant communication and in the case of identified risk, the Wall Guides can identify the location of the user and direct them to use local crisis services. Contact information for local crisis services was provided on the unique landing page created for the Ontario users. The BWW components include (1) educational material, (2) guided support courses based on principles of cognitive behavioral therapy and behavior change, and (3) text communication posts as either 1:1 with another member or a Wall Guide or open to discussion groups composed of any members of the peer community. The educational materials and guided support courses cover a wide range of mental health related topics including grief, depression, anxiety, smoking cessation, substance use, trauma, among others. Users can also post and comment on *bricks*—creative self-expressions whereby the user designs a brick which they can place in the digital *wall*. See Figures 1-4 for screenshots of the BWW components. The BWW reaches out to users through their registration email if there has been a prolonged period of inactivity, encouraging users to log on. In this study, all participants received an email alias that linked to their personal email and a unique prescription for a BWW account. A follow-up call was made by a study team member 3 days after sending the prescription and again after 2 weeks, if the account had not been activated. Technical support was available from the research team, the BWW, and the Ontario Telemedicine Network.

XSL•FO
RenderX

Figure 1. The Big White Wall offers self assessments across a range of mental health concerns including depression, anxiety, and substance use.

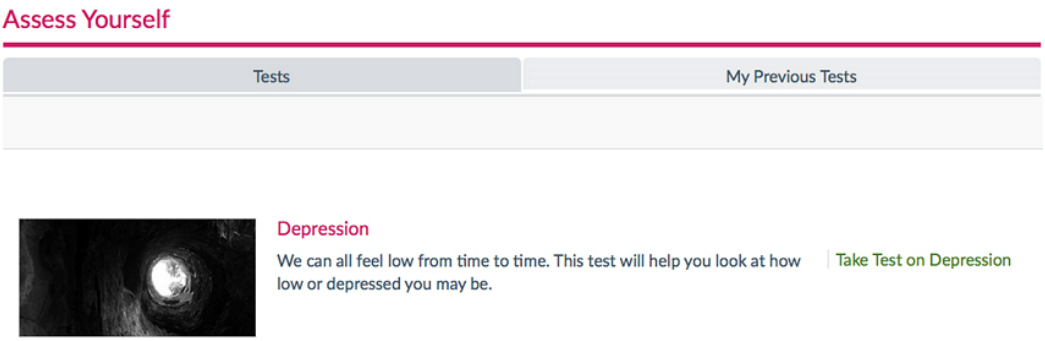


Figure 2. The Useful Stuff pages provide information on mental health conditions and interventions.

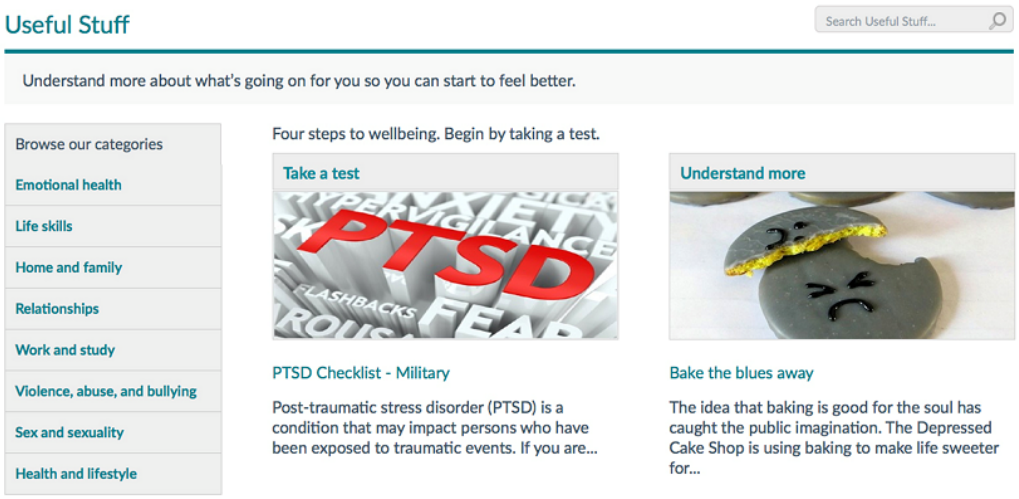
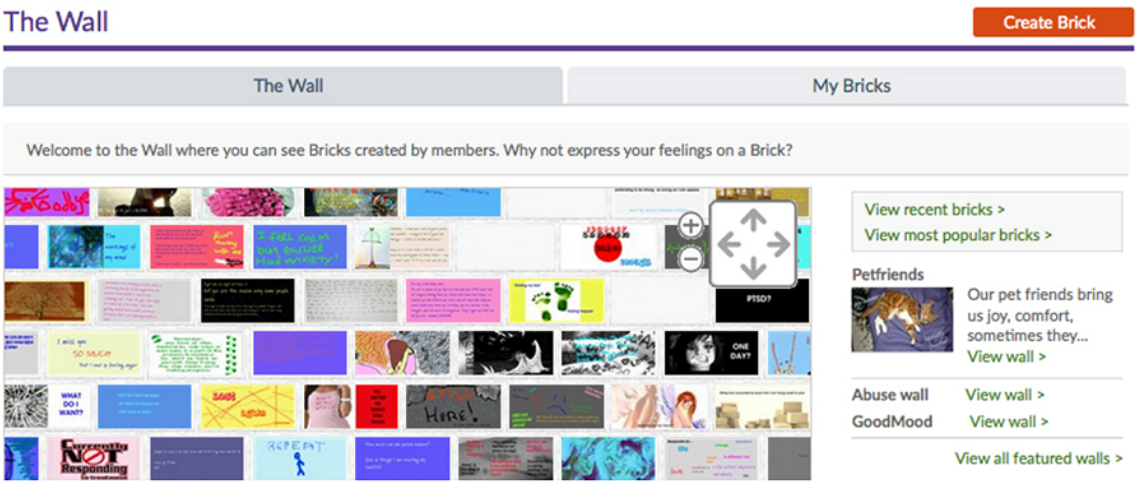
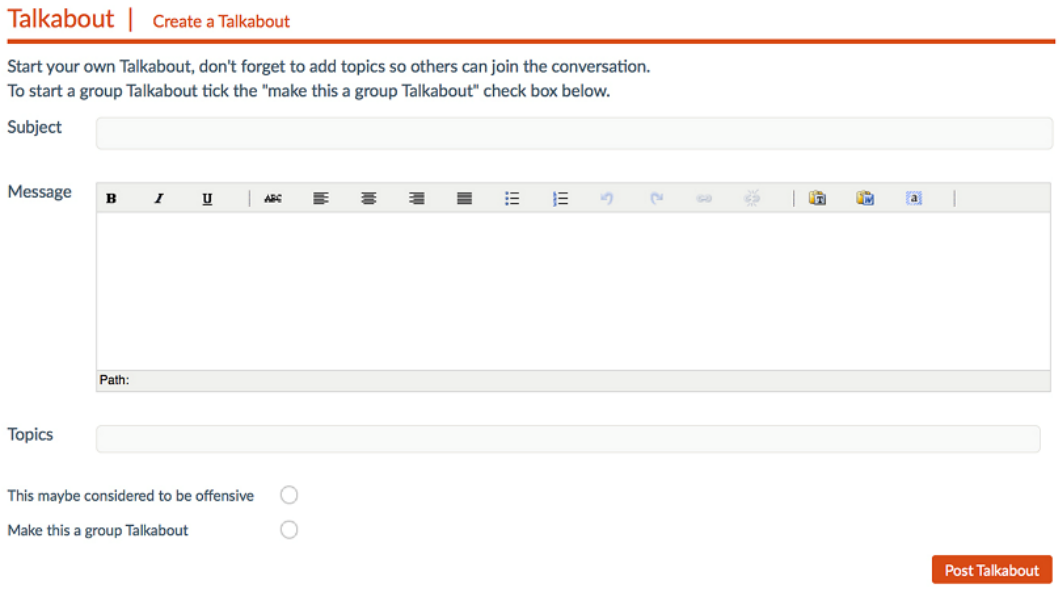


Figure 3. The Wall is a space for users to post self-expression statements through the creation of artistic bricks. Users can also comment on each other's bricks.



XSL•FO
RenderX

Figure 4. Moderated personal and group Talkabouts allow users to converse with wall guides and peers regarding their mental health concerns and experiences.



Follow-Up

Follow-up data were collected between October 2016 and April 2017. All follow-up data were collected by self-report via electronic surveys through REDCap or collected by phone or in person by a study team member and subsequently input into the REDCap database.

Participants received an automated survey link by email 1 week before 3 months postrandomization with a follow-up personalized email containing the survey link. In these emails, DTG participants were reminded that they would receive access to the BWW once the survey was completed, although all participants were ultimately given access regardless of survey completion. At 3 months and 3 months plus 1 week, reminder phone calls were made. Surveys were closed 2 weeks after the 3-month time point. Study team members collecting the outcome assessments were blinded to group allocation. Both baseline and 3-month postrandomization surveys were completed via the Web-based survey in 95% of cases, with the remainder by phone or in person.

Outcomes

The primary outcome was mental health recovery at 3 months assessed with the Recovery Assessment Scale-revised (RAS-r). This outcome assesses an individual's orientation toward recovery and self-management across 5 domains: (1) personal confidence and hope, (2) willingness to ask for help, (3) goal and success orientation, (4) reliance on others, and (5) not dominated by symptoms [20]. The study intervention claims to promote self-management specifically, and the RAS-r was chosen as an outcome relevant to participants across all diagnoses. The use of the RAS-r also reflects the current *recovery era* for mental health policy and services with the focus of treatment shifting to consumers finding satisfying and fulfilling lives, rather than being symptom free [20]. The RAS-r is the most widely used and validated recovery assessment tool and has been studied in several patient populations and shown to correlate with symptoms and function [20]. It is a 24-item scale, with all items scored on a 5-point scale from *strongly disagree* to *strongly agree*, with total scores ranging from 24 to 120.

Secondary outcomes were the PHQ-9 to assess symptoms of depression, Generalized Anxiety Disorder Questionnaire-7 item (GAD-7) to assess symptoms of anxiety, EuroQOL 5-dimension quality of life questionnaire (EQ-5D-5L) to assess quality of life, and the Community Integration Questionnaire (CIQ) [21]. The PHQ-9 scale contains 9 items rated on a 4-point Likert scale from *never* to *almost every day*, with a higher score representing more symptoms [22]. The GAD-7 scale has 7 items rated on a 4-point Likert scale from *never* to *almost every day*, with a higher score representing a higher likelihood of an anxiety disorder. The EQ-5D-5L questionnaire from the EuroQOL group [23] comprises 5 dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety/depression, which are rated on 5 levels and can be interpreted as an index score by comparing with available normative values for adult populations. The EQ-5D-5L is paired with a visual analog scale (VAS) rated from 0 to 100 to assess perceived overall health at the time of survey completion. The CIQ consists of 15 items and is intended as a brief, reliable measure of a person's level of integration into the home and community. The overall score can range from 0 to 29 with a higher score indicating better integration [24].

Actual BWW utilization data were obtained for all ITG participants including activation of account and number of logins.

Sample Size

The target sample size for our trial was set by the sponsors at 1000 participants. We aimed to recruit this number of participants, with an expectation that loss to follow-up would be approximately 30% based on data from other similar trials [25]. We calculated the minimal detectable difference between the 2 treatment groups for a linear regression analysis controlling

XSL•FO
RenderX

for baseline score assuming a 0·8 correlation between baseline and 3-month follow-up RAS-r measurements. Assuming 30% loss to follow-up, for a sample size of 700 after attrition, allocated in a 2:1 ratio, using an alpha of .05 and power of 0.9, the minimal detectable difference was 1.35 on the RAS-r.

Randomization, Concealment, and Blinding

Participants were randomized 2:1 to the ITG and DTG groups. A 2:1 randomization ratio was used to offer the intervention immediately to a higher number of participants and increase recruitment. Randomization sequences, using block sizes of 3 or 6 were computer generated by an organization external to the research team, with stratification by site and recruitment setting. Group allocation sequence was concealed but once allocated, participants were not blinded. Participants allocated to the DTG group received a telephone or email notification that they would gain access to the website after a 3-month delay. Most follow-up data were collected by Web-based survey, but in the case where follow-up data were collected by an assessor, the assessor was not the same as the person who did the initial data collection, as this person also disclosed randomization to the participant and would not be blinded at follow-up. Data analysts remained blinded throughout the study.

Statistical Analysis

The primary outcome, RAS-r, was analyzed at 3 months for all participants regardless of whether they activated their BWW account, but individuals with missing data at 3 months were excluded. We planned to repeat the analysis using a marginal structural model to account for attrition; however, owing to higher than expected attrition that was nonrandom, this analysis was not completed. RAS-r scores at 3 months were modelled with a linear regression controlling for baseline RAS-r score and treatment group. A second analysis controlled for prespecified covariates including baseline PHQ-9, baseline GAD-7, age, gender, recruitment setting, and age of first onset of mental health problems. The same analysis was repeated for all secondary outcomes at 3 months. Analyses were completed after data collection was complete. BWW utilization data were analyzed descriptively to illustrate uptake of the intervention in the ITG participants only.

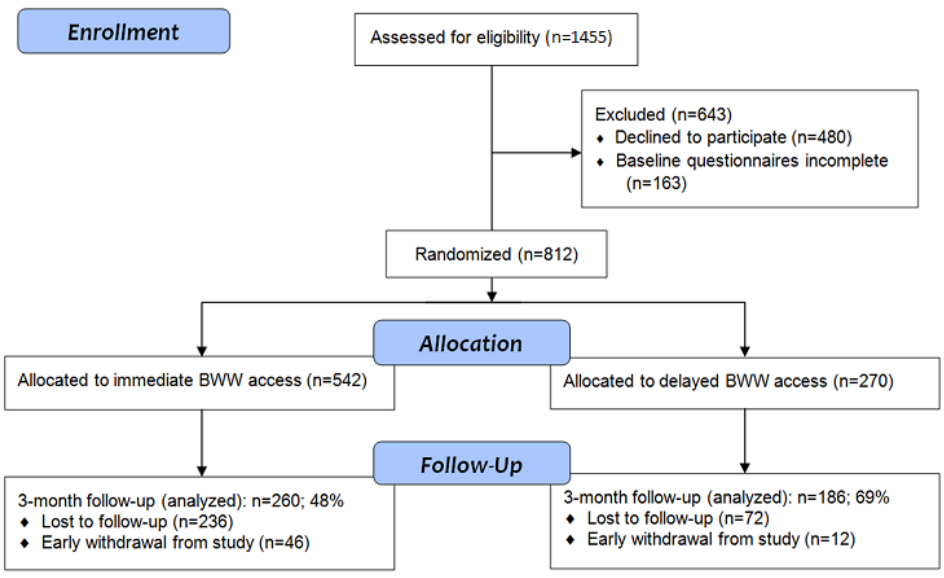
Exploratory Analysis

Among ITG participants who completed the 3-month outcome measures, we examined whether there was an association between actual utilization of the application and the primary and secondary outcomes. Utilization of the application was defined by number of BWW logins, with users categorized into 4 groups based on distribution quartiles: 0 to 1 login, 2 to 3 logins, 4 to 9 logins, and 10 or more logins. Owing to degree of missingness, primary and secondary outcomes were separately modeled with repeated measures ANOVA using time as a repeated measure. We tested for an interaction between BWW use and time, after adjusting for same covariates as in the main analysis. Significant main effects were explored with post-hoc Bonferroni tests adjusted for multiple comparisons.

Results

We approached 1455 individuals, of whom 975 (67.0%) consented. Of the 975, 163 (16.7%) did not complete the baseline assessment questionnaire, leaving 812 to be randomized, 270 (33.3%) to DTG and 542 (67.7%) to ITG (see Figure 5). Distribution of participants across recruitment settings was as follows: adult mood and anxiety psychiatry programs (n=294/812, 36.2%), youth mood and anxiety programs (n=42/812, 5%), mood and anxiety psychotherapy programs (n=73/812, 9%), emergency department/urgent care clinic (n=139/812, 17.1%), borderline personality/trauma therapy programs (n=114/812, 14.0%), and substance use program (n=150/812, 18.5%). Follow-up was achieved for 446/812 (54.9%) at 3 months, 260/542 (48.0%) in the ITG group and 186/270 (68.9%) in the DTG. At baseline, the randomized groups were well-balanced in terms of sociodemographic, mental health variables, and previous 3-month health care utilization (Table 1).

Figure 5. CONSORT flow diagram of participants through the trial. BWW: Big White Wall.



XSL•FO
RenderX

Table 1. Baseline characteristics of study participants by group.

Variables	Immediate Treatment Group (n=542)[a]	Delayed Treatment Group (n=270)[a]
Age (years), mean (SD)	41.5 (13.4)	40.0 (13.9)
Gender, n (%)		
Male	143 (27)	59 (22)
Female	391 (73)	207 (78)
Transgendered	1 (0)	1 (0)
Ethnicity, n (%)		
White	440 (82)	219 (83)
Non-white	98 (17)	46 (18)
Relationship status, n (%)		
In a relationship	286 (53)	127 (47)
Not in a relationship	252 (47)	142 (52)
Employment status, n (%)		
Full-time (including homemaker with young children)	181 (33)	96 (35)
Part-time/volunteer/homemaker without young children	100 (19)	41 (15)
Not working (retired owing to age or actively looking for work)	33 (13)	26 (14)
Not working (not looking for work)	84 (35)	57 (31)
Household income in Can $, n (%)		
<35K	238 (48)	126 (52)
35K-50K	56 (11)	38 (16)
50K-80K	82 (16)	25 (10)
>80K	123 (25)	52 (22)
Age first experienced mental health problems (years), mean (SD)	18.7 (12.5)	19.0 (12.7)
Age first sought help (years), mean (SD)	26.7 (12.9)	26.3 (13.0)
Taking medication at baseline, n (%)	438 (81)	204 (76)
Previous 3-month hospitalization, n (%)	47 (9)	23 (9)
Previous 3-month emergency room visit, n (%)	77 (14)	27 (10)
Agree with: Self-help tools helpful for people with mental health problems, n (%)		
Definitely agree	203 (38)	110 (41)
Somewhat agree	310 (58)	151 (56)
Somewhat or completely disagree	24 (4)	8 (3)
How much expected improvement in mental health through BWW[b,c], n (%)		
Less than 50%	194 (38)	110 (42)
50%	141 (26)	76 (29)
More than 50%	179 (35)	76 (29)

[a]Percentages calculated after missing data removed.

[b]Responses were recorded in 10% increments but based on their distribution have been recategorized.

[c]BWW: Big White Wall.

There were some differences between those who were lost to follow-up and those who completed the 3-month survey overall and in each group (see Multimedia Appendices 1 and 2). Overall, the survey completers were older, and more likely to be working full-time or unemployed and not looking for work. ITG survey completers were more likely to activate their BWW accounts and have more logins than noncompleters in that group, with the same age and employment patterns as overall. In the DTG group, the only significant difference found was for recruitment setting; those from youth and emergency or urgent care settings were most likely to complete follow-up. A proportion of survey noncompleters from both the ITG and

XSL•FO
RenderX

DTG groups withdrew early from the study (46 out of 282 ITG survey noncompleters (16%) and 12 out of 84 (14%) DTG survey noncompleters).

Primary Outcome

The primary analysis showed a statistically significant difference between the ITG and DTG groups for the RAS-r at 3 months, with the ITG participants having an RAS-r score on average 5.34 points higher than the DTG participants at 3 months (Table 2). In the adjusted model, the effect remained significant but slightly lower in magnitude (Table 2).

Table 2. Baseline and 3-month primary and secondary outcomes among survey completers with linear regression results for outcomes at 3 months.

Outcome	Immediate treatment group (n=260), mean (SD)		Delayed treatment group (n=186), mean (SD)		Unadjusted treatment effect size (95% CI)[a]	Adjusted treatment effect size (95% CI)[a,b]
	Baseline	3 months	Baseline	3 months		
Primary outcome						
Recovery Assessment Scale, revised	77.4 (14.0)	83.3 (15.1)	76.3 (14.1)	77.2 (14.6)	5.32 (3.33 to 7.31)	4.97 (2.90 to 7.05)
Secondary outcomes						
Patient Health Questionnaire-9 item	14.8 (6.9)	11.5 (6.4)	16.0 (6.5)	14.2 (6.8)	−1.95 (−2.94 to −0.95)	−1.83 (−2·85 to −0·82)
Generalized Anxiety Disorder Questionnaire-7 item	11·5 (5·6)	9·1 (5·3)	12·3 (5·6)	11·4 (5·7)	−1.75 (−2.60 to −0.90)	−1.56 (−2.42 to −0.70)
EuroQOL 5-dimension quality of life questionnaire	0.68 (0.16)	0.71 (0.17)	0.68 (0.16)	0.69 (0.15)	0.014 (−0.009 to 0.036)	0.008 (−0.015 to 0.031)
EuroQOL Visual Analog Scale	56.8 (19.2)	58.8 (21.5)	55.1 (19.8)	55.4 (21.9)	2.55 (−1.17 to 6.26)	1.68 (−1.97 to 5.32)
Community Integration Questionnaire	16.9 (5.0)	17.0 (5.2)	16.7 (4.6)	17.1 (4.8)	−0.32 (−0.93 to 0.29)	−0.30 (−0.93 to 0.33)

[a]Delayed treatment group is the reference group in all analyses; all models include baseline score.

[b]Adjusted for age, sex, recruitment setting, baseline PHQ-9 score, baseline GAD-7 score, and age of first onset of mental health problems.

Secondary Outcomes

PHQ-9 and GAD-7 scores were significantly lower in the ITG group compared with the DTG group in the main and adjusted analyses (Table 2). No statistically significant differences were found for EQ-5D-5L index score, EQ-VAS, or CIQ (Table 2).

Participant Safety

The data safety monitoring board reviewed data at 3 and 6 months after the start of recruitment and did not find any substantial increase in suicidal ideation between groups to warrant investigation or early termination. Over the course of the study, one death was reported to the study team and deemed unrelated to the intervention.

Uptake, Use, and Satisfaction With the Intervention

Big White Wall Activation

Among the 542 participants who were randomized to receive immediate access to the BWW, 76 (14%) never activated their BWW account during the study period. Half of these individuals (n=39, 51%) withdrew early from the study. Reasons for early withdrawal included the following: technical issues, loss of perceived need or interest, and lack of time owing to competing priorities.

Utilization

There was large variability in the range of total number of times that participants logged on to the site, from 0 to 236 times. The mean number of logins was 8.7, with a standard deviation of 18.1, a median of 2, and a mode of 1. Only 58% (312/542) of participants logged on 2 or more times, with approximately 20% of all participants accounting for 80% of the total logins.

Exploratory Analysis

There was no significant interaction between BWW use and time for the primary outcome of RAS-r. The interaction was significant for PHQ-9 ($F_{3,257}$=4.14; P=.007) and GAD-7 ($F_{3,267}$=3.89; P=.009). In post-hoc analysis for PHQ-9, a significantly greater reduction in score over time was present for the groups with 10 or more logins and 2 to 3 logins, relative to the 0 to 1 login group (4.14 points vs 1.43, P=.03 and 5.00 vs 1.43, P=.02, respectively). In post-hoc analysis for GAD-7, only the group with 2 to 3 logins had a significantly greater reduction in score over time compared with the group with 0 to 1 login (4.36 points vs .91 points, P=.006). There was no significant effect of the interaction between BWW use and time on the other secondary outcomes.

Discussion

Principal Findings

Immediate access to the BWW resulted in small, significant improvements in mental health recovery, as well as depressive and anxiety symptoms at 3 months compared with those randomized to delayed access. These statically significant findings are limited by high, differential drop out between treatment groups and overall, and were below minimal clinically important differences for these outcome measures (eg, a

XSL•FO
RenderX

difference of 1.8 on the PHQ-9 where a clinically important change is 5 points) [22]. Engagement with the intervention varied highly and interestingly, we observed the commonly found Pareto principle for population effects—also known as the *80/20* rule [26]. That is, 20% of users accounted for approximately 80% of the activity. We found some evidence of a user effect whereby those who completed follow-up and engaged with the application were more likely to experience improvements in symptoms of depression and anxiety. This was not linear, however, with the user group having 2 to 3 logins experiencing the most consistent improvements relative to the lowest user group. To our knowledge, this is the first randomized evaluation of the BWW, and one of a few large multicenter pragmatic trials of any multicomponent internet-based mental health intervention. The study population represents more treatment-refractory and severely symptomatic individuals than are usually included in studies of Web-based mental health applications.

Comparison With Other Studies

Although the BWW is unique in comparison with other studied applications, the observed low engagement is similar [25]. A comparatively large pragmatic trial that examined 2 computerized modular CBT programs with added telephone support reported that fewer than 20% of participants had completed all modules at follow-up [25]. Modular-style courses are only one component of the BWW, the other components being peer support, artistic self-expression, and more general psychoeducation. Compared with a modular program, it is more difficult to define an adequate *dose* for multicomponent interventions such as the BWW [13]. We examined different self-directed *doses* of the BWW based on logins and did not find a linear relationship between higher use and better outcomes, although it appeared that some engagement may have been better than none. Some users may benefit from a few targeted uses to get direction or motivation. A systematic review examining the relationship between e-therapy adherence and outcomes reported that studies targeting depression did not find a significant impact of total number of logins on outcomes in contrast to studies examining physical health outcomes like weight management or smoking [9]. Studies using modular interventions, however, did report that module completion led to better outcomes. Conversely, one of the main features of the BWW used more often by the higher engagers is its moderated virtual community of peers [15], an intervention for which the evidence has yielded some very mixed results [27,28]. Some authors have described a potential harm through aggravated symptoms and what Takahashi et al termed the *downward depressive spiral* which was linked to higher depressive tendencies at baseline [29]. This may have tempered the improvements in the higher engager groups or conversely may have led to early disengagement from the application as a result of negative reactions to the peer community.

Limitations

This study experienced loss to follow-up that was proportionally higher in the intervention group. High drop out of up to 50% in trials of Web-based mental health interventions has been reported [10] and retention in digitized trials has specifically been discussed as a challenge [30]. In our study, the issues encountered with our automated survey were addressed quickly but may have impacted survey completion. More specifically, other studies of Web-based mental health programs have also reported disproportionate dropout in the intervention group [10]. This trend likely represents some study disengagement for intervention-specific reasons such as dislike of the intervention format or lack of perceived utility. In this study, some very early technology issues with BWW activation likely affected early user engagement and the Web-based platform did not include a mobile version of the site or application which could have deterred use for some. To partially address this limitation, we conducted the exploratory analysis on study completers.

We chose the pan-diagnostic RAS-r as the primary outcome, given that our study sample was recruited from a range of mental health settings and participants had a range of mental health needs. The promise of internet-based mental health interventions specifically to support recovery has been discussed [31]; however, this outcome lacks established clinically important cutoffs. Secondary outcomes may have been affected by the lack of diagnostic specificity relative to other studies that have focused on specific diagnoses such as depression and anxiety [6]. In this study, we evaluated the BWW as a solution for any mental health or addiction-related need; although we adjusted for recruitment program, it is possible that the BWW may work better for certain diagnostic or need subgroups, which represents an area for further evaluation.

Conclusions and Policy Implications

From this trial, we cannot definitively conclude the effectiveness of this or similar solutions at a population level for individuals accessing specialized mental health care. However, internet-based interventions built on evidence-based principles of mental health care are likely to be beneficial where users are motivated to engage and where the format of the application is a good fit. Determining this subset and/or how to effectively motivate more people to engage with the interventions is a critical next step. Pragmatic studies that build in process evaluations and subset analyses that examine moderators of use and related outcomes are required. This study was undertaken as part of a large provincial implementation program with multiple research, clinical, and policy stakeholders involved throughout, a type of integrated, real-world approach to evaluation that is essential to establish effective health systems solutions [32].

Acknowledgments

The research team would like to acknowledge the site leads, project coordinators, and research assistants for their roles in project planning, recruitment, and data collection: Beth Brannon, Anusha Baskaran, and Alifa Siddiqui (Ontario Shores Centre for Mental

XSL•FO
RenderX

Health Sciences); Paul McGary, Tom McHugh, and Joanne Gourgouvelis (Lakeridge Health); Joanne Leung (Women's College Hospital). This study was sponsored by the Ontario Telemedicine Network through funding from Canada Health Infoway and the Ontario Ministry of Health and Long Term Care.

Authors' Contributions

All authors contributed to the conceptualization of the study and development of the study protocol. Data analyses for all trial outcomes were completed by AC. Additional analyses on BWW utilization data were completed by MN, NO, and JMH. The manuscript was drafted by JMH and revised critically by all authors. All authors have approved the submitted version of the manuscript. JMH acts as the guarantor for this work and affirms that the manuscript is an honest, accurate, and transparent account of the study being reported. Authors had full access to all study data (including statistical reports and tables) and take responsibility for the integrity of the data and the accuracy of the data analysis. Participant-level data are available from the corresponding author upon request.

Conflicts of Interest

PA consults with and receives personal fees from the study sponsor on projects unrelated to this study. All other authors declare no competing interests that could influence the submitted work. The sponsor was engaged in the selection of the technology vendor, conducted the implementation of the intervention at the sites, provided input into the study design, and funded the study completion. The sponsor was not involved in the data analysis or writing of the manuscript. The sponsor reviewed the manuscript for accuracy or reporting related to the sponsor involvement. The vendor (BWW) was not involved in the study design, development, implementation, analysis, or writing of this manuscript. The researchers acted independently of all study sponsors and the vendor.

Multimedia Appendix 1

Baseline characteristics of study participants who did and didn't complete the follow-up measures at 3 months.

[PDF File (Adobe PDF File), 90KB - jmir_v21i6e10838_app1.pdf]

Multimedia Appendix 2

Baseline characteristics of study participants by group, who did and didn't complete the follow-up measures at 3 months.

[PDF File (Adobe PDF File), 63KB - jmir_v21i5e10838_app2.pdf]

Multimedia Appendix 3

CONSORT-eHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - jmir_v21i6e10838_app3.pdf]

References

1. Whiteford HA, Ferrari AJ, Degenhardt L, Feigin V, Vos T. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PLoS One 2015;10(2):e0116820 [FREE Full text] [doi: 10.1371/journal.pone.0116820] [Medline: 25658103]
2. Fleury M, Grenier G, Bamvita J, Perreault M, Kestens Y, Caron J. Comprehensive determinants of health service utilisation for mental health reasons in a Canadian catchment area. Int J Equity Health 2012 Apr 2;11:20 [FREE Full text] [doi: 10.1186/1475-9276-11-20] [Medline: 22469459]
3. Mojtabai R, Olfson M, Sampson NA, Jin R, Druss B, Wang PS, et al. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. Psychol Med 2011 Aug;41(8):1751-1761 [FREE Full text] [doi: 10.1017/S0033291710002291] [Medline: 21134315]
4. Vasiliadis H, Lesage A, Adair C, Boyer R. Service use for mental health reasons: cross-provincial differences in rates, determinants, and equity of access. Can J Psychiatry 2005 Sep;50(10):614-619. [doi: 10.1177/070674370505001007] [Medline: 16276852]
5. Marzano L, Bardill A, Fields B, Herd K, Veale D, Grey N, et al. The application of mHealth to mental health: opportunities and challenges. Lancet Psychiatry 2015 Oct;2(10):942-948. [doi: 10.1016/S2215-0366(15)00268-0] [Medline: 26462228]
6. Cunningham JA, Gulliver A, Farrer L, Bennett K, Carron-Arthur B. Internet interventions for mental health and addictions: current findings and future directions. Curr Psychiatry Rep 2014 Dec;16(12):521. [doi: 10.1007/s11920-014-0521-5] [Medline: 25308390]
7. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]

XSL•FO
RenderX

8. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
9. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
10. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
11. Greenhalgh T, Shaw S, Wherton J, Hughes G, Lynch J, A'Court C, et al. SCALS: a fourth-generation study of assisted living technologies in their organisational, social, political and policy context. BMJ Open 2016 Feb 15;6(2):e010208 [FREE Full text] [doi: 10.1136/bmjopen-2015-010208] [Medline: 26880671]
12. Batterham PJ, Sunderland M, Calear AL, Davey CG, Christensen H, Teesson M, et al. Developing a roadmap for the translation of e-mental health services for depression. Aust N Z J Psychiatry 2015 Apr 23:776-784. [doi: 10.1177/0004867415582054] [Medline: 25907269]
13. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
14. Naylor D, Girard F, Mintz J, Fraser N, Jenkins T, Power C. Health - Canada.ca. 2018. Unleashing innovationxcellent healthcare for Canada - report of the advisory panel on healthcare innovation URL: https://tinyurl.com/qx2cf8z [accessed 2019-05-13] [WebCite Cache ID 78LLlkldd]
15. Harding C, Chung H. Behavioral health support and online peer communities: international experiences. Mhealth 2016;2:43 [FREE Full text] [doi: 10.21037/mhealth.2016.10.04] [Medline: 28293613]
16. Hensel JM, Shaw J, Jeffs L, Ivers NM, Desveaux L, Cohen A, et al. A pragmatic randomized control trial and realist evaluation on the implementation and effectiveness of an internet application to support self-management among individuals seeking specialized mental health care: a study protocol. BMC Psychiatry 2016 Oct 18;16(1):350 [FREE Full text] [doi: 10.1186/s12888-016-1057-5] [Medline: 27756281]
17. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [Medline: 11132119]
18. Price M, Anderson P, Henrich CC, Rothbaum BO. Greater expectations: using hierarchical linear modeling to examine expectancy for treatment outcome as a predictor of treatment response. Behav Ther 2008 Dec;39(4):398-405 [FREE Full text] [doi: 10.1016/j.beth.2007.12.002] [Medline: 19027436]
19. 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]
20. Salzer MS, Brusilovskiy E. Advancing recovery science: reliability and validity properties of the Recovery Assessment Scale. Psychiatr Serv 2014 Apr 1;65(4):442-453. [doi: 10.1176/appi.ps.201300089] [Medline: 24487405]
21. Willer B, Ottenbacher KJ, Coad ML. The community integration questionnaire. A comparative examination. Am J Phys Med Rehabil 1994 Apr;73(2):103-111. [Medline: 8148099]
22. 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]
23. Euroqol. 2018. Valuation of EQ-5D URL: https://euroqol.org/eq-5d-instruments/valuation-of-eq-5d/ [accessed 2019-05-14] [WebCite Cache ID 78NX02aCi]
24. Stålnacke B. Post-traumatic stress, depression, and community integration a long time after whiplash injury. Ment Illn 2010 Jan 25;2(1):e4 [FREE Full text] [doi: 10.4081/mi.2010.e4] [Medline: 25478087]
25. Gilbody S, Littlewood E, Hewitt C, Brierley G, Tharmanathan P, Araya R, et al. Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. Br Med J 2015;351:h5627 [FREE Full text] [Medline: 26559241]
26. Robertson TM, Lofgren RP. Where population health misses the mark: breaking the 80/20 rule. Acad Med 2015 Mar;90(3):277-278. [doi: 10.1097/ACM.0000000000000627] [Medline: 25551859]
27. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): do ISGs reduce depressive symptoms? J Med Internet Res 2009 Sep 30;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
28. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. Br Med J 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
29. Takahashi Y, Uchida C, Miyaki K, Sakai M, Shimbo T, Nakayama T. Potential benefits and harms of a peer support social network service on the internet for people with depressive tendencies: qualitative content analysis and social network analysis. J Med Internet Res 2009 Jul 23;11(3):e29 [FREE Full text] [doi: 10.2196/jmir.1142] [Medline: 19632979]

XSL•FO
RenderX

30. 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]
31. Strand M, Gammon D, Ruland CM. Transitions from biomedical to recovery-oriented practices in mental health: a scoping review to explore the role of Internet-based interventions. BMC Health Serv Res 2017 Dec 7;17(1):257 [FREE Full text] [doi: 10.1186/s12913-017-2176-5] [Medline: 28388907]
32. Barratt H, Shaw J, Simpson L, Bhatia S, Fulop N. Health services research: building capacity to meet the needs of the health care system. J Health Serv Res Policy 2017 Dec;22(4):243-249 [FREE Full text] [doi: 10.1177/1355819617714815] [Medline: 28786700]

Abbreviations

BWW: Big White Wall
CIQ: Community Integration Questionnaire
DTG: delayed treatment group
EQ-5D-5L: EuroQOL 5-dimension quality of life questionnaire
GAD-7: Generalized Anxiety Disorder Questionnaire-7 item
ITG: immediate treatment group
PHQ-9: Patient Health Questionnaire-9 item
RAS-r: Recovery Assessment Scale, revised
VAS: visual analog scale

Edited by N Cahill, S Hagens, G Eysenbach; submitted 20.04.18; peer-reviewed by M Ashford, B Sakakibara, W Martinez; comments to author 22.03.19; revised version received 26.04.19; accepted 26.04.19; published 04.06.19.

Please cite as:
Hensel JM, Shaw J, Ivers NM, Desveaux L, Vigod SN, Cohen A, Onabajo N, Agarwal P, Mukerji G, Yang R, Nguyen M, Bouck Z, Wong I, Jeffs L, Jamieson T, Bhatia RS
A Web-Based Mental Health Platform for Individuals Seeking Specialized Mental Health Care Services: Multicenter Pragmatic Randomized Controlled Trial
J Med Internet Res 2019;21(6):e10838
URL: https://www.jmir.org/2019/6/e10838/
doi:10.2196/10838
PMID:31165710

©Jennifer M Hensel, James Shaw, Noah M Ivers, Laura Desveaux, Simone N Vigod, Ashley Cohen, Nike Onabajo, Payal Agarwal, Geetha Mukerji, Rebecca Yang, Megan Nguyen, Zachary Bouck, Ivy Wong, Lianne Jeffs, Trevor Jamieson, R Sacha Bhatia. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Valuing Citizen Access to Digital Health Services: Applied Value-Based Outcomes in the Canadian Context and Tools for Modernizing Health Systems

Christina Hackett[1], PhD; Kelsey Brennan[1], MA; Heather Smith Fowler[1], MA; Chad Leaver[2], MSc, MBA

[1]Social Research and Demonstration Corporation, Ottawa, ON, Canada

[2]Canada Health Infoway, Toronto, ON, Canada

Corresponding Author:
Christina Hackett, PhD
Social Research and Demonstration Corporation
55 Murray St, Suite 400
Ottawa, ON, K1N5M3
Canada
Phone: 1 6132374312
Fax: 1 6132375045
Email: chackett@srdc.org

Abstract

Background: In publicly funded health systems, digital health technologies are strategies that aim to improve the quality and safety of health care service delivery and enhance patient experiences and outcomes. In Canada, governments and health organizations have invested in digital health technologies such as personal health records (PHRs) and other electronic service functionalities and innovation across provincial and territorial health systems.

Objective: Patients' access to their own information via secure, Web-based PHRs and integrated virtual care services are promising mechanisms for supporting patient engagement in health care. We draw on current evidence to develop an economic model that estimates the demonstrated and potential value of these digital health initiatives.

Methods: We first synthesized results from a variety of Canadian and international studies on the outcomes for patients and service providers associated with PHRs across a continuum of services, ranging from viewing information (eg, laboratory results) on the Web to electronic prescription renewal to email or video conferencing with care teams and providers. We then developed a quantitative model of estimated value, grounded in these demonstrated benefits and citizen use (2016-2017). In addition to estimating the costs saved from patient and system perspectives, we used a novel application of a compensating differential approach to assess the value (independent of costs) to society of improved health and well-being resulting from PHR use.

Results: Patients' access to a range of digital PHR functions generated value for Canadians and health systems by increasing health system productivity, and improving access to and quality of health care provided. As opportunities increased to interact and engage with health care providers via PHR functions, the marginal value generated by utilization of PHR functionalities also increased. Web-based prescription renewal generated the largest share of the total current value from the patient perspective. From the health systems perspective, Canadians' ability to view their information on the Web was the largest value share. If PHRs were to be implemented with more integrated virtual care services, the value generated from populations with chronic illnesses such as severe and persistent mental illness and diabetes could amount to between Can $800 million and Can $1 billion per year across Canadian health systems.

Conclusions: PHRs with higher interactivity could yield substantial potential value from wider implementation in Canada and increased adoption rates in certain target groups—namely, high-frequency health system users and their caregivers. Further research is needed to tie PHR use to health outcomes across PHR functions, care settings, and patient populations.

(J Med Internet Res 2019;21(6):e12277) doi:10.2196/12277

KEYWORDS

personal health records; patient portals; electronic health records; patient engagement; health care costs; cost of illness; economic evaluation

XSL•FO
RenderX

Introduction

Health Systems and Digital Health Technologies

Digital health technology is an umbrella term encompassing a variety of innovations such as mobile health, health information technology, wearable devices, telehealth and telemedicine, and personalized medicine. Digital health technologies are envisioned as a step forward in empowering patients to participate in their health care decisions and, thereby, *democratize* health care. In publicly funded health systems, digital health technologies are strategies aimed at improving the quality and safety of health care service delivery as well as enhancing patient experiences and outcomes [1-3]. For example, the United Kingdom and Australia have launched large-scale personal health records (PHRs) initiatives—the NHS.UK and My Health Record, respectively—to connect patients with their electronic health and medical records [4-6].

In Canada, federal and provincial governments and health organizations have invested in digital health technologies and innovation across provincial and territorial health systems, including PHRs with a variety of electronic service (e-service) functionalities [7]. Canada Health Infoway is a federally funded, nonprofit organization whose mandate is to "realize the vision of healthier Canadians through innovative digital health solutions." Through Infoway, Canada has invested in the development, implementation, and evaluation of various digital health solutions, including PHRs.

Personal Health Records and Health System Improvement

A PHR is defined as a digital space, or Web/mobile, application-based, health or medical record that holds all or a portion of clinical information about an individual (eg, laboratory results or prescribed medications) [8-11]. Multiple terms exist to refer to platforms through which patients can access their health care information, including PHR, patient portal, or consumer portal (to maintain consistency, we use PHR throughout). PHRs can be either stand-alone and driven by information entered by individuals and stored on their personal devices or tethered to a health care organization or system. Information shared via PHRs can come from multiple sources and is managed, shared, and controlled by the individual patient; therefore, PHRs should conform to nationally recognized interoperability standards [12].

PHRs in the scope of analyses presented in this paper:

1. are integrated with point-of-care clinical health information systems, allowing individuals to view their health record information from electronic medical records, hospital (or hospital consortium) information systems, laboratory testing and results information systems, electronic health records (EHRs), or health information exchanges;
2. are tethered to a health care organization or system, and are linked to information provided within the EHR attached to that system (hospital, insurance plan, or other health care organization); and
3. contain connected care functionalities/services such as communication and consultation features with point-of-care clinics, providers, or care organizations.

Stand-alone PHRs were considered out of scope, including goal- or outcome-defined clinician-patient coaching models that allow patients to input information and PHRs that do not link to clinical-source information systems.

Patients access their PHRs through a secure personal identification authorization process. PHRs can provide a range of e-service functionalities, categorized in-depth by the Center for Information Technology Leadership as comprising 4 overarching types of informational interaction—collection, sharing, self-management, and exchange [13]. *How* PHRs facilitate these interactions can vary; functions include the ability to (1) view health care information (electronic view [e-View]), (2) exchange secure emails and messages with health care providers (electronic visit [e-Visit]), (3) renew prescription medication on the Web (electronic prescription renew [e-Rx renew]), and (4) visit virtually via videoconference with their health care providers (virtual visit). These functionalities are illustrated in Figure 1, alongside Canadian adoption and use rates from 2016 and 2017. As such, PHRs have the potential to influence value-based outcomes in terms of avoided direct and indirect costs but also as a mechanism to improve longer-term trends in health literacy, patient satisfaction [14], access and user-centered design [15-17], and clinical outcomes.

The terms p activation, engagement, empowerment, and patient-centered care, although used and measured distinctly, are often used interchangeably and broadly refer to mechanisms by which patients and caregivers interact with their care trajectory and the degree to which they feel confident and satisfied in doing so [18]. Efforts to enhance patient activation and patient engagement in health and health care, as well as improve patient experience, are increasingly at the forefront of health policy and strategy in both high-resource and low- to middle-income health systems worldwide [19-21].

A growing body of literature highlights how PHRs can increase patient activation by providing seamless access points through which patients access their health care information and communicate with their providers or care teams [11,22,23]. Patients' interaction with their health care information has been shown to positively influence patient engagement, health behaviors, and associated health outcomes in a variety of settings [24-26]. However, this body of work has focused largely on patient activation in health care and health outcomes and less on the *process* by which activation occurs through engagement with digital health solutions [27-29]. For the remainder of the paper, we use the term *patient activation* to refer to the desired outcome of patients' interactions with their PHR and *patient engagement* as the mechanism through which PHRs facilitate patient activation.

XSL•FO
RenderX

Figure 1. Definition and utilization of electronic service functionalities in Canada (2016-2017).

Electronic view

Electronic visit

Virtual visit

Electronic prescription renew

Citizen access to their health information	Secure e-mail with health provider/place of care	Secure face-to-face video visit with health provider	Request for prescription renewal
Definitions			
A personal health record includes the primary function of viewing health information: **e-view**	An **e-visit** is a patient service that allows patients/ caregivers the ability to communicate with their health care team through secure email or text messaging	A **virtual visit** is a patient service that allows patients/caregivers to meet with their health care team virtually face-to-face, through functions such as video calls	**e-Rx renew** is a patient service that allows patients/caregivers to renew prescriptions
Percentage of Canadians who reported using in the last year			
Accessed their medical records online [7%*-8%]	Consulted with health care providers online via e-mail [5%-8%]	Visited virtually with provider [3%*-4%]	Sent a web-based prescription renewal request [10%-12%]

A note on the current adoption and use of personal health records to support the quantitative model:

We used the percentage of Canadians who reported they accessed their medical record online or used specific e-service functionalities in 2016 (minimum) and 2017 (maximum), to estimate the current value related to citizen access and use of personal health records and connected care e-services in Canada. The quantitative model also adjusted estimates accordingly for e-view and e-Rx renew functionalities by the proportion of Canadians who said they had a lab test or filled a prescription in the past year.

*Here adoption rates are rounded to the nearest integer – precise Canadian and provincial adoption rates are available in Multimedia Appendices 1 and 2.

PHRs have been shown to influence patient engagement by reducing barriers to access to information via seamless Web-based platforms [18]. In certain high-volume health management organizations in the United States (eg, Kaiser Permanente and Veterans Affairs), where tethered PHRs have been implemented and value-based outcomes evaluated, PHR functionalities involving high levels of patient interaction with their regular health care providers (via e-Visits and virtual visits) have been shown to improve health outcomes, while reducing per capita costs of care [21,41]. Despite these benefits, it is important to note that a digital divide exists between those who can and do access these technologies and those for whom there are barriers to using technology to engage with their health care provider. Differences in access are influenced by multiple sociodemographic factors. However, there is little consensus about determinants of access and uptake across intersections of individual-level characteristics: often there are differences in use by socioeconomic status, education, and age, those in lower-income quintiles having lower levels of education and those of older age tend to access digital health technologies less [30-32]. Physical access to internet and to technological devices, in addition to technology-related skills and literacy, all contribute to perpetuating or bridging the digital divide in high-resource contexts [33,34]. A recent electronic health equity framework notes the complex interplay of macro (socio-techno-economic-political context), patient (social position and intersection between various demographics), intermediary (material circumstances and social capital), and digital technology implementation factors that influence access to health technologies [35].

Evaluation of Value-Based Outcomes and Digital Health Technology/Personal Health Records

Although there are a growing number of economic evaluations of digital health solutions, there are conceptual, methodological, and practical challenges to assessing the economic benefit of implementing and adopting PHRs, both in terms of increased health system efficiency and improved quality of and access to medical care [10,11]. These challenges include the lack of a clearly defined perspective resulting from multiple investors and sources of funding, the lack of identified options for comparison, and having comparable costs and outcomes across different PHR interventions/sites [11,36]. In addition, economic analyses and evaluations often take place in single health care settings with a focus on a specific patient population, and are not easily interpretable at a systems level. Conversely, analyses focusing solely on the benefits of PHRs and associated cost-savings or financial outcomes for a broader population are

XSL•FO
RenderX

often based on relatively few sources of evidence, which may or may not reflect the health system contexts in which the benefit estimates are applied.

Using an established approach developed by DeLone and McLean [37] for evaluating information technology initiatives, Canada Health Infoway (Infoway) has developed a comprehensive benefits evaluation framework for sites implementing PHRs, inclusive of net benefits or value-based outcomes [38]. This framework facilitates the evaluation of implementation factors as well as aligning outcomes with PHR use.

Our objectives in this study were as follows:

1. To synthesize outcomes generated by benefits evaluations conducted at multiple sites implementing PHRs in Canada, across different types of care settings, and serving different patient populations.
2. To estimate the relative economic benefit of implementing PHRs in those different care settings and patient populations compared with business as usual (either before or in the absence of PHR implementation) from 3 perspectives—health system (payer), patient and caregiver, and the economic benefit to society resulting from improved population health (societal perspective).

We used contextually specific data as well as cost and outcome data from the peer-reviewed literature to demonstrate the current and potential added value generated by PHR adoption and implementation at a national level across various health care settings, patient populations, and 4 specific PHR functionalities. In addition to estimating value based on resources saved by patients and caregivers and clinician and clinic productivity gains, we used a novel application of various approaches to empirically link PHR use and improved health, health behaviors, and increased life satisfaction.

Methods

Overview

We developed 3 quantitative models to estimate the economic benefit of patients and health care organizations having access to PHRs, relative to what occurred or would occur in the absence of being able to access PHRs (ie, business as usual). Typically, economic analyses of interventions or technologies in health care estimate the comparative effectiveness, utility, or broader opportunity cost between 2 or more alternative interventions [11,39]. Given the variability of implementation costs, relative provincial investment in digital health initiatives, and amount of financial support allocated to each PHR implementation site in Canada, we opted to present a cost-outcome description that synthesizes the aggregate population-level economic benefits of PHR adoption and implementation in Canada. Specifically, we applied a 1-sided economic analysis that estimates the economic benefits of the relative health care quality, access, and productivity gains from PHR implementation versus nonimplementation. Each of the 3 models reflects a different perspective from which we conducted our analyses:

1. Patient and caregiver perspective
2. Health systems perspective
3. Societal perspective/improved health

Data and Model Inputs

We drew on the benefits evaluation studies of recent PHR initiatives conducted by Canadian health care organizations (n=12) to derive estimates of any added economic benefits. Data from these evaluations represented PHR sites in 5 provinces (British Columbia, Saskatchewan, Ontario, Quebec, and Nova Scotia). Each study employed a system and use survey that examined the use of PHR functions as well as PHR user experiences and also explored quantitative outcomes of use. In addition, the sites used administrative records and user reports to assess the odds of user versus nonuser requests for health care information, average number of visits avoided (for both groups), and resulting organizational efficiency gains (eg, full-time equivalent staffing saved). Further information regarding the studies included in our 3 quantitative models can be found in Multimedia Appendices 1 and 2.

To expand our models to include relevant benefits found in the peer-reviewed literature, we conducted a systematic search of the following electronic bibliographic databases: EconLit, Health Systems Evidence, PsycINFO, and MEDLINE. Our key search terms included personal health record OR patient portal AND (economic benefit OR benefit). The following limits were applied to all searches: (1) keyword search to abstract and title, (2) English or French, and (3) 2010 to 2017. A team of 2 researchers conducted the searches and subsequent stages of review and extraction.

Selection of Model Inputs

For inclusion in our study, the peer-reviewed and gray literature had to (1) provide estimates of the economic benefit of PHR use from the patient, system, or societal perspectives, (2) provide clinical benefits of PHR use, and (3) take place in a high-resource health system context. The excluded studies were those without a counterfactual scenario (ie, PHR use compared with no health system use), those that included PHRs with functionalities out of scope, or intervention studies without a quantified clinical outcome clearly attributable to PHR use. After applying the selection criteria, a total of 81 studies were selected for full-text review, and 21 of these were used as the primary input sources to the quantitative models developed.

From studies selected for inclusion (including the benefits evaluations conducted in Canada), 2 researchers reviewed each document and extracted data on clinical and economic benefits of PHR use. A third member of the research team (CL) validated each extracted benefit and model input. The outcomes reported by respondents in these evaluations generated a range of values of effectiveness, using a series of classification factors in the Infoway Benefits Evaluation framework. As seen in Table 1, these classification fields included PHR functionality, care setting, benefit domain, and benefit recipient or patient population targeted by the intervention and care setting.

XSL•FO
RenderX

Table 1. Classification factors for benefit estimation.

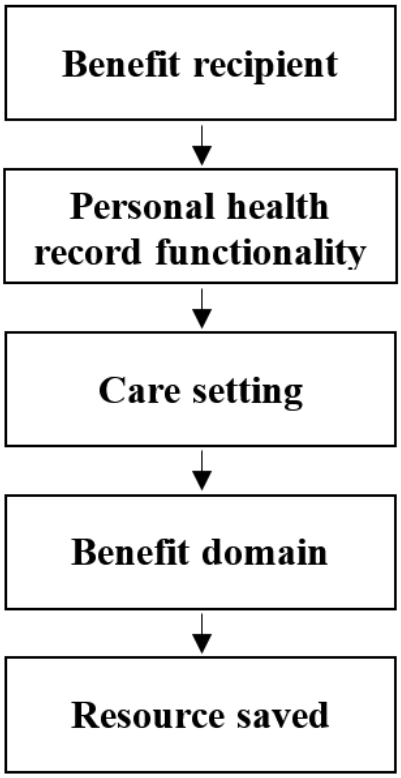
Field of classification	Definition
PHR[a] functionality	**The method by which the PHR engages patients/caregivers.**
Electronic view	Primary function of PHR is the viewing of health information.
Electronic visit	A patient e-service that allows patients and their caregivers the ability to communicate with their health care team through secure email or short message service text messaging.
Virtual visit	A patient e-service that allows patients and their caregivers the ability to meet with their health care provider via a face-to-face virtual encounter through functions such as video calls.
Electronic prescription renew	A patient e-service that allows patients and their caregivers to renew prescriptions.
Care setting	**The medical care setting in which the benefit of PHR was found to accrue.**
Primary care	Day-to-day health care delivered by a health care provider (eg, general practitioner's office).
Specialist care—mental health	Health care provided for issues related to mental health, including community-based and inpatient care.
Specialist care—chronic conditions	Health care provided for issues related to other chronic conditions such as diabetes.
Hospital-based care	Inpatient and outpatient care provided in hospital/hospital-affiliated settings.
Pediatric care	Health care provided to children.
Benefit domain	**Areas of value.**
Quality	An increase in health quality as a result of PHR use, such as increased healthy behaviors, improved health outcomes, or increased life satisfaction.
Productivity	An increase in productivity as a result of PHR use, such as saved time or resources.
Access	An increase in access to health care as a result of PHR use.
Benefit recipient	**The recipient of the benefit.**
Patient/caregiver	The benefit accrued directly to the patient or caregiver.
Health system	The benefit accrued to the health system (eg, the primary care provider or the hospital).
Health outcomes	The benefit involved an improvement in population health.
Resource saved	**The way in which patients/caregivers, the health system, and health outcomes benefitted from PHR use.**
Avoided visits to health care providers	The primary way by which patients/caregivers benefitted from PHR use, including saved time and cost related to travel, saved time arranging caregiving and caregiving costs, and avoided time off work.
Increased productivity among health systems	The primary way by which health systems benefitted from PHR use, including avoided visits and reduced calls from patients/caregivers, avoided emergency department visits, and avoided preventable adverse drug events.
Improved healthy behaviors	A way in which health outcomes benefitted from PHR use, such as better medication adherence.
Increased life satisfaction	A way in which health outcomes benefitted from PHR use, using a validated life satisfaction scale.
Improved health	A way in which health outcomes benefitted from PHR use, through changes in patient activation.

[a]PHR: personal health record.

Once we classified areas of value to benefit domains, as outlined above, we created a hierarchy of specific value-based outcomes, as outlined in Figure 2, such that benefit recipients were the primary classification factor, followed by the associated PHR functionality, care setting, benefit domain, and empirical value of resource saved. Estimates were adjusted to the specific yearly (2016 and 2017) utilization rates for each PHR functionality, the Canadian population aged over 18 years reporting access or use of indicated health services, and affected clinical populations for health outcomes estimates, both nationally and by care setting.

XSL•FO
RenderX

Figure 2. Hierarchy of benefit classification.



As it is unknown if and to what extent certain PHR functionalities might increase health care costs, we defined and included value-based outcomes in terms of the change in units of resources used (by patients and health systems) relative to a counterfactual scenario. The extracted outcomes data represented the added value of PHR use reported compared with an alternative scenario or counterfactual. For example, 1 health system outcome was *avoided medical errors* achieved through patient use of an electronic prescription renewal when compared with traditional prescription renewal processes. The value to the health system in this regard results from reduced preventable adverse events (primarily drug-related events) and a resulting reduction in unnecessary health system utilization. To avoid double-counting benefits or discounting costs related to PHR implementation in terms of patient and health system resources, we included only benefits that were clearly separate from resource costs. For example, because e-Visit functionalities in primary care settings involve physician time, we did not count avoided in-person visits to primary care physicians as an economic benefit to the health system but did for patients who saved time and financial costs.

Multimedia Appendices 3, 4, and 5 provide an overview of the included Infoway studies, the final sample of respondents from which the reported outcomes are derived, the counterfactual/comparison group or alternative scenario used for comparison, and the type of indicators used for the patient, health system, and societal perspectives, respectively. In addition to the outcomes extracted from Canadian PHR initiatives, we drew on various sources to identify relevant estimates for classification factors from publicly available data. See Multimedia Appendices 1 and 2 for an overview of the data sources used to estimate economic benefits Canada-wide from patient/caregiver and health system cost perspectives, respectively.

Model Estimation

We defined the economic benefit or value *V* to each perspective as the reported costs to patients and health systems in the absence of PHRs, less the cost savings *S* reported as a result of PHR adoption or implementation within a health care setting. We offset the cost savings reported by subtracting an estimate of the deadweight, or cost savings that may have occurred naturally, in the absence of PHR adoption.

$$V_{Patient} = C_{PCurrent} - (S_T + S_{LF} + S_C - D)$$

Costs to patients and caregivers were expenses related to attending an in-person appointment with their health care provider. Resources saved by patients included travel costs and travel time, caregiving costs and time spent arranging care, and reduced time away from paid employment. Direct average costs were obtained from primary data collected from patients' and caregivers' self-reporting from Canadian PHR implementations and used as model inputs. In terms of indirect costs, the amount of time saved was obtained from patients' and caregivers' self-reporting and valued using median Canadian income for time away from paid employment. For time costs that did not relate to paid employment, time was valued at 25% and 50% of median income. Multimedia Appendix 1 contains details related to the sources used to value patient/caregiver costs, and Multimedia Appendix 6 provides a narrative summary of how we calculated estimates from this perspective.

$$V_{Health\ system} = C_{HSCurrent} - (S_{U(AV,H0)} + S_{PG} + S_{PS} - D)$$

From the health system perspective, current costs are a function of the quantity of health care services provided within the

XSL•FO
RenderX

relevant health care setting and to the relevant patient population and the cost per service. Resources saved for health systems include health care provider time (reduction in clinician time to complete tasks/patient consultation), increased productivity, and savings resulting from improved patient safety (eg, reduced preventable adverse drug events). For example, patients and caregivers having access to prescription information on the Web can increase patient and caregiver awareness of their current and previous medications, resulting in decreased prescription error on the supply side (pharmacists and health care providers) and increased medication adherence by patients. See Multimedia Appendices 2 and 7 for details regarding valuation sources and an overview of methods used to estimate the value of citizens' use of PHR functionalities to health systems in Canada.

$$V_{Population} = C_{HSCurrent} - (V_{Satisfaction} + S_{Health} + S_{Health\ behavior} - D)$$

Where there are tangible costs avoided to the health system attributable to PHR use (eg, unnecessary in-person visits avoided and provider time on operational tasks mitigated by Web-based communication or automation), estimates of benefit to the health care systems in Canada are more straightforward than for less tangible benefits such as improved health, health care behaviors, and health status. To provide estimates of improvements to behaviors and health status, we converted positive health outcomes reported in the amassed evidence base into specific econometric indicators of resources saved, according to associated PHR functionalities.

In addition to estimating the value of improved health behavior and health status of PHR users compared with nonusers, we created a direct empirical link between improvement in life satisfaction reported by adults with severe and persistent mental illness as a result of using PHR functionalities and a household income equivalent. We applied a compensating differential approach developed by Helliwell and Huang (2010) [40] that takes the difference in life satisfaction before/after for PHR users versus nonusers and then divides this by the coefficient for the marginal effect of life satisfaction on household income for Canadians (beta=0.14). For a detailed overview of the approaches used to estimate the value of improved health behaviors and health status, see Multimedia Appendix 8.

Range of Values

In each of the 3 models, we estimated a range of value for benefits currently realized as well as the potential value for these initial benefit areas with advanced utilization rates. Ranges in current utilization rates of available PHR functionalities were measured in 2016 and 2017 from nationally representative surveys designed by Canada Health Infoway and administered by an independent research organization [7,8]. To model potential value-based outcomes, we applied projected utilization rates of 25%, 35%, and 50%. The magnitude of ranges illustrated in our final estimates represent variations in the reported outcomes realized, in the way in which personal time can be valued as a proportion of income, and in reported uses of the health care system.

Assumptions

The assumptions underlying our calculations were as follows:

1. The benefits realized by the study sample from source evidence reflect benefits that could be realized by populations with similar case mixes or diagnoses.
2. Benefits realized by health care organizations in terms of increased productivity are transferable to other similar contexts across Canada.
3. The proportion of the population represented by each province in Canada can be used to expand province-level benefits to national-level benefits. Conversely, national-level benefits can be disaggregated into those at the provincial-level.
4. Outcomes related to PHR effectiveness remain constant over time.

All the estimates generated by valuing provider time (ie, time saved by health care providers and organizations) were created on the basis of a range of values from the least expensive unit of time (eg, an administrative staff) to the most expensive unit of time (eg, a physician).

Results

Overview

Table 2 summarizes the outcomes identified in our review by PHR functionality and outlines areas in which PHRs have led to identified benefits and in which care settings those benefits were realized. Current Canadian PHR adoption rates range from 3% to 4% for virtual visit technology to 10% to 12% for e-Rx renew functionalities. The benefits realized included avoided visits and time costs saved across different care settings, such as community-based clinics, primary care clinics, and hospitals.

XSL•FO
RenderX

Table 2. Summary of evidence of Canadians' current use of personal health records and electronic services.

Personal health record functionality	Electronic view	Electronic visit	Virtual visit	Electronic prescription renew
Adoption rate	7%-8% of Canadians can and have accessed their health care information on the Web.	5%-8% of Canadians can and have communicated with their health care provider securely on the Web.	3%-4% of Canadians can and have visited virtually with their health care provider securely on the Web.	10%-12% of Canadians can and have renewed their prescription medication on the Web.
Care setting	Primary care, hospital care, and community-based mental health services were care settings where Canadians and health systems benefited from accessing their health care information on the Web.	Community-based mental health services were the care setting where Canadians and health systems benefited from communicating with their health care provider securely on the Web.	Primary care was the care setting where Canadians and health systems currently benefited from visiting virtually with their health care provider on the Web. Primary care networks for people with chronic conditions demonstrated potential benefits to Canadians and health systems.	Community-based mental health services and hospitals were the care settings where Canadians and healthy systems benefited from renewing their prescription medication on the Web.
Resource saved (patient)	Canadians benefited by avoiding visits to primary and mental health care providers.	Canadians benefited by avoiding visits to primary and mental health care providers.	Canadians benefited by avoiding visits to primary health care providers.	Canadians benefited by avoiding visits to primary and mental health care providers.
Resource saved (health system)	Health systems benefited by increased productivity (time saved because of avoided visits and calls).	Health systems benefited by increased productivity (resources saved because of avoided emergency department visits).	Health systems benefited by increased productivity (resources saved because of avoided emergency department visits). Potential benefits were also identified through increased access via remote care provision.	Health systems benefited by increased productivity (time saved because of avoided visits) and increased quality (preventable adverse drug events avoided).

Patient/Caregiver Perspective

Current Benefit

For patients and caregivers, direct value from PHRs currently manifests in the form of avoided in-person visits to health care providers, leading to savings in travel time and costs, caregiving costs and time arranging care, and time off work. These savings are generally shared across e-View, e-Visit, e-Rx renew, and virtual visit functionalities, with variations resulting from differing adoption rates, care settings, and the number of Canadians accessing certain types of health care services (ie, Canadians with a primary care physician or Canadians having a laboratory test in the past year).

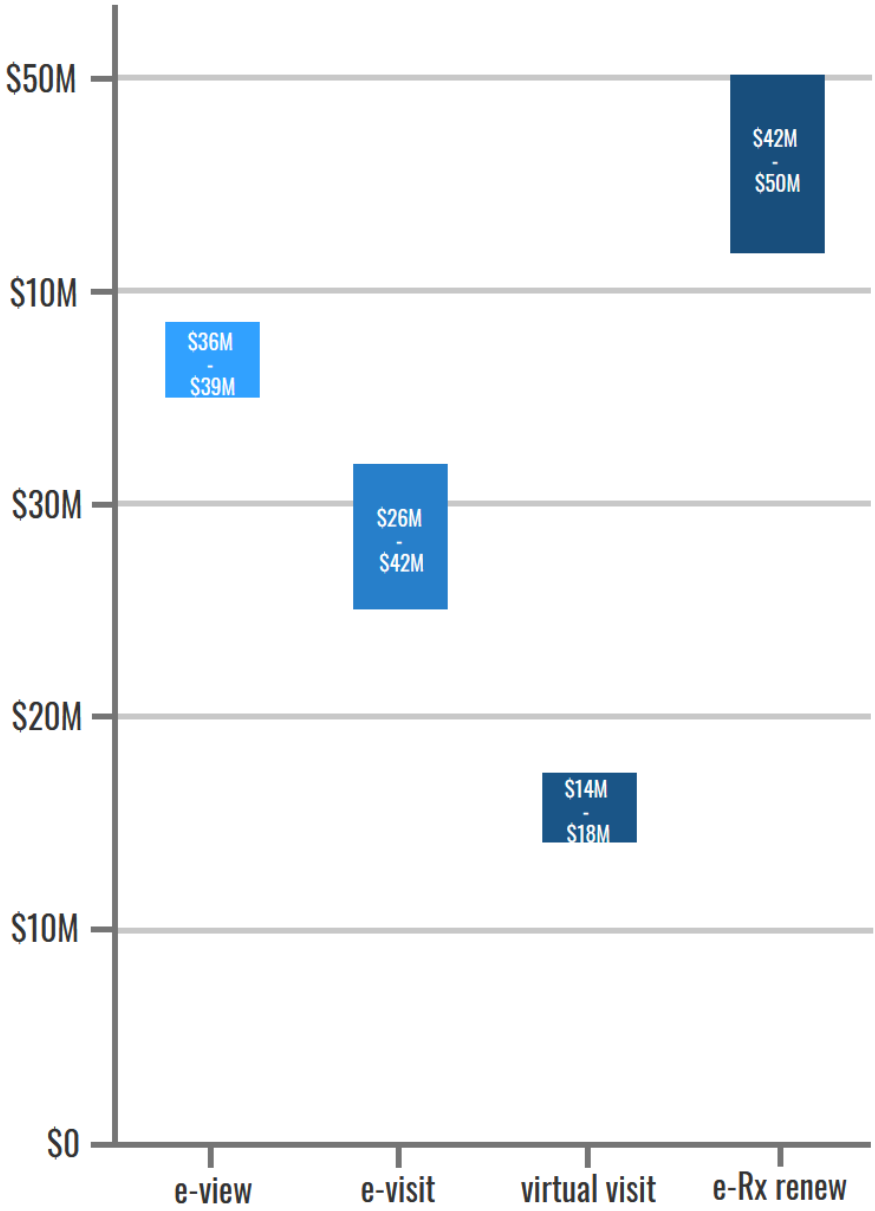
Evidence of the value of PHRs for patients and caregivers in Canada has so far been demonstrated in primary care and mental health care settings (see Multimedia Appendix 9). Given this, model estimates were associated with population-level access estimates in these care settings. The value to Canadians who use currently available PHR functionalities is estimated to be Can $119 to Can $150 million per year (aggregated across all Canadians at current adoption rates). Figure 3 demonstrates the way in which this value is distributed across PHR functionalities: overall, our patient model demonstrates that most of the current value (2016-2017) to patients and caregivers is derived from e-Rx-renew (34%), followed by e-Visit (28%), e-View (26%), and virtual visit (12%).

Canadians have varying health care needs and, accordingly, different patterns of service use within the health care system. These factors affect the range and magnitude of value to citizens from use of the PHR functionalities estimated in this study and are important to interpret in context. We present below the range of total value for Canadian population subgroups based on frequency of visits per year [8], represented in terms of avoided costs because of avoided unnecessary in-person visits, weighted by utilization rates of PHR functionalities and the proportion of the population that reported each frequency of visit to their health care provider. The current annual aggregated value to Canadian patients/caregivers based on health system utilization for low-, medium-, and high-volume health system user segments of the Canadian population is as follows:

1. Can $61 to 153 million for the 41% of the Canadian population that makes 1 to 2 visits to a health care provider per year,
2. Can $87 to 148 million for the 19% of the Canadian population that makes 3 to 4 visits to a health care provider per year, and
3. Can $143 to 180 million or more for the 19% of the Canadian population that makes 5 or more visits to a health care provider per year (value estimates are based on 5 visits per year).

XSL•FO
RenderX

Figure 3. Current annual value to Canadians who use personal health records, by functionality. e-View: electronic view; e-Visit: electronic visit; e-Rx renew: electronic prescription renew.

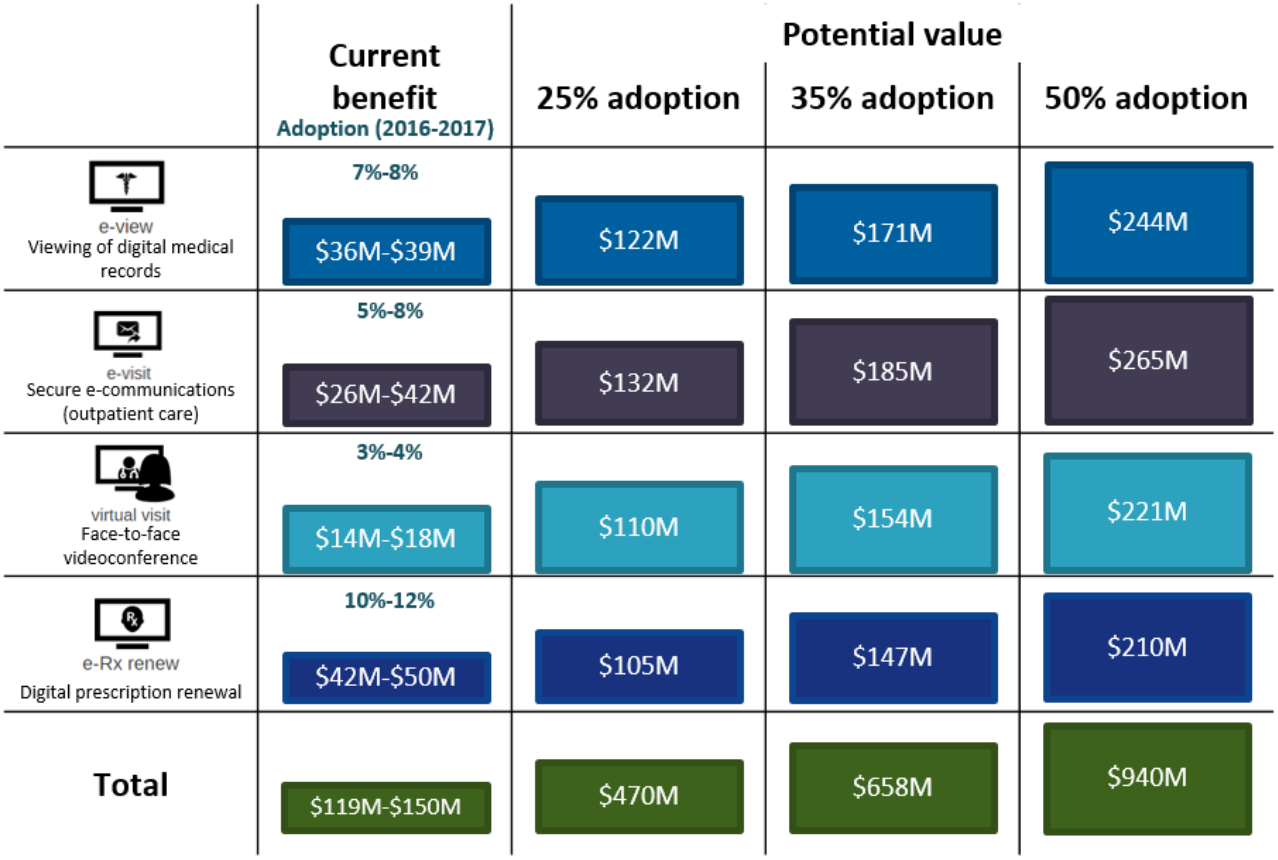


Potential Benefit

On the basis of these value estimates, if Canadians were to increase PHR adoption from current adoption rates to 25%, 35%, or 50%, the total value to patients and caregivers would increase to Can $470 million, Can $658 million, and Can $940 million, respectively. Figure 4 represents how increasing adoption rates of each PHR functionality would increase the value from the patient perspective, assuming a steady-state of benefits realized (ie, that the benefits to Canadians remained constant over time).

XSL•FO
RenderX

Figure 4. Projection of patient/caregiver benefits with increased personal health record adoption. e-view: electronic view; e-visit: electronic visit; e-Rx renew: electronic prescription renew.



Health System Perspective

Current Benefit

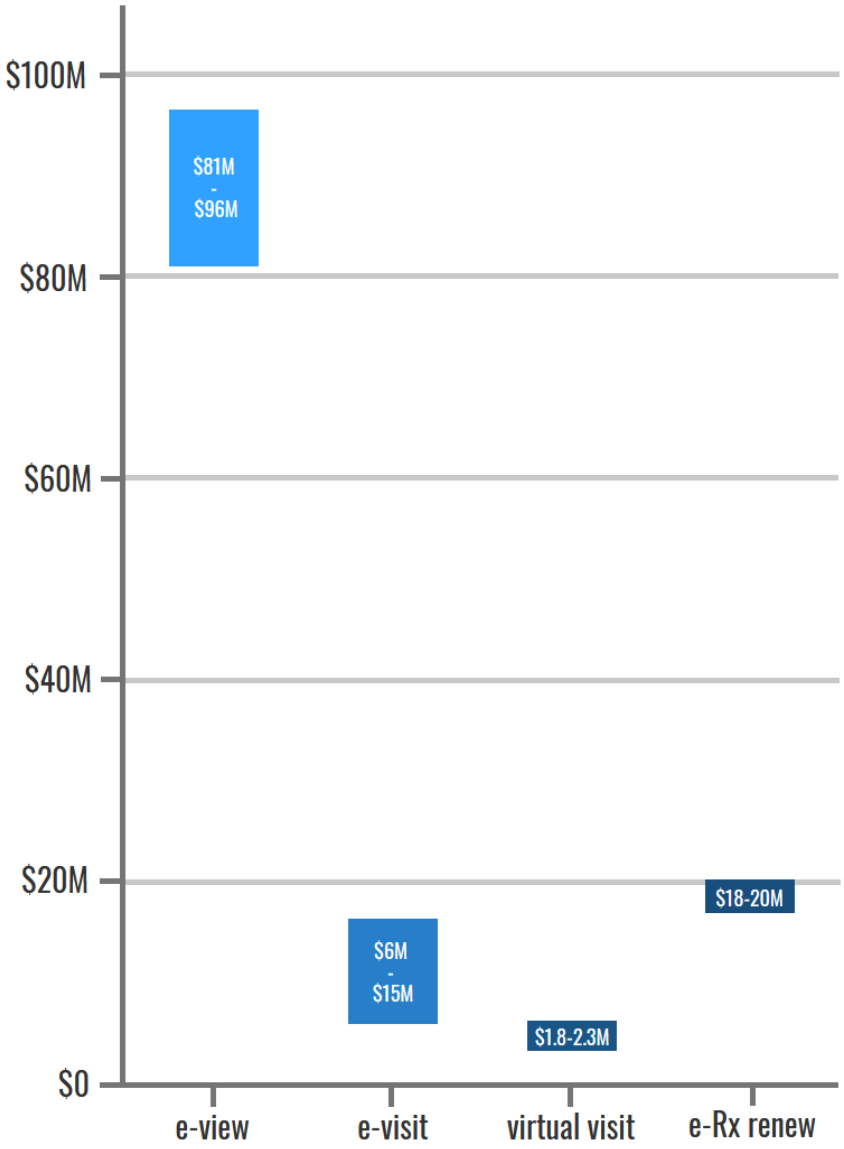
From the health system perspective, we found that direct value from PHRs currently manifests in the form of increased productivity, leading to clinic or clinician time saved because of avoided visits and operational and clinician processes involved in responding to and resolving patient phone calls. There is also some evidence of increased health care quality, manifested by avoided preventable adverse drug events. Evidence of the value of PHRs for health systems in Canada has so far been demonstrated in primary care, hospital, pediatric, and mental health care settings, with value also identified internationally in other areas of specialized care (eg, veteran health services and chronic illnesses; see Multimedia Appendix 10). The non-Canadian evidence mentioned above also shows promising evidence of benefits for virtual visit functionalities.

The current estimated value to health systems where PHR functionalities are in use is estimated to be Can $106 to Can $134 million a year. Figure 5 demonstrates the way in which that value is distributed across PHR functionalities. Overall, the health system model demonstrates value primarily related to e-View (72%)—the most widely available and utilized functionality—followed by e-Rx renew (15.2%), e-Visit (11.1%), and virtual visit (1.7%). We see below that most of the cost savings to the health system are driven by e-View functionality and are estimated at between Can $81 and Can $96 million per year.

Under the assumption that the savings associated with virtual visit use for specialized care can be realized in Canada at the same rate it has been realized internationally, savings for that functionality would rise from Can $27 to Can $54 million. In addition, in 2 of the identified studies, outcomes of PHR use were measured separately for different functionalities, allowing the model to explore the marginal benefit of increased PHR-enabled provider interaction within the same patient population. Using these studies, the model for valuing health system benefits of PHR use found that with increased opportunities for patients to directly consult or communicate with their providers, the value to the health system increased by 24% to 32%. In other words, moving from being able to passively view health care information on the Web to communicating or consulting virtually with a health care provider generated substantially more value to both patients and the health system. For example, in 1 study [31], the estimated benefit of patients using only e-Rx renew in a hospital setting was Can $105 million, whereas the estimated benefit of patients who used the e-Visit function was Can $127 million.

XSL•FO
RenderX

Figure 5. Current annual value to health systems where personal health record functionalities are in use. e-view: electronic view; e-visit: electronic visit; e-Rx renew: electronic prescription renew.

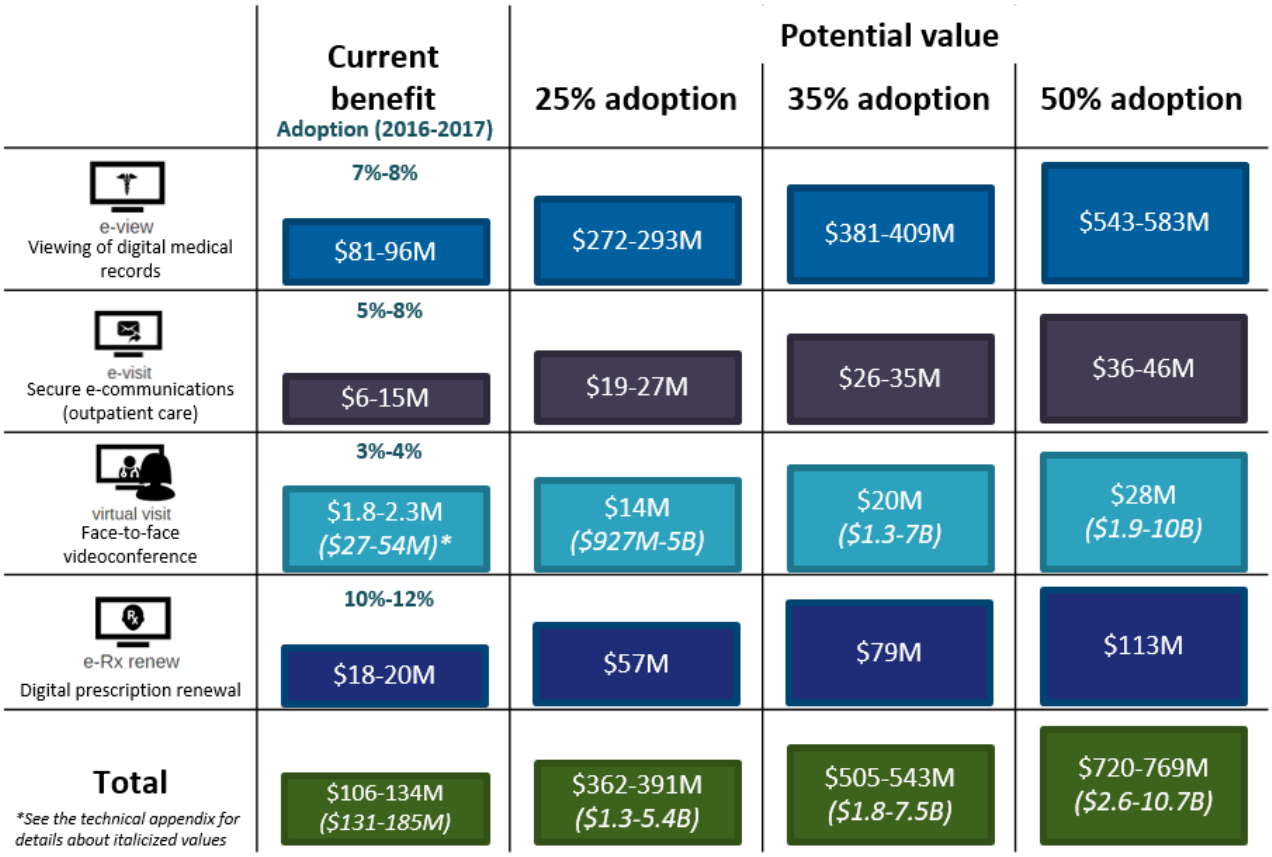


Potential Benefit

Using the estimates of value to health systems in Canada, we found that if Canadians were to increase PHR adoption to 25%, 35%, or 50%, the total value to patients and caregivers would increase to Can $362 to Can $391 million, Can $505 to Can $543 million, and Can $720 to Can $769 million, respectively. Figure 6 represents how increasing adoption rates of each PHR functionality would increase the value from the health system perspective, assuming a steady state of benefits realized. The second set of totals (represented by italicized figure ranges in parentheses below) represent the integration of benefits estimates that assume increased adoption and integrate evidence from jurisdictions outside of Canada with wrap-around PHR functionalities (inclusive of virtual visits). If Canadians had access to comparable PHRs, the total value to health systems in Canada would be Can $131 to Can $185 million at current adoption rates, and when projected to 25%, 35%, and 50%, the estimated values would range from Can $1.3 to Can $5.4 billion, Can $1.8 to Can $7.5 billion, and Can $2.6 to Can $10.7 billion, respectively.

XSL•FO
RenderX

Figure 6. Projection of health system benefits with increased personal health record adoption. e-view: electronic view; e-visit: electronic visit; e-Rx renew: electronic prescription renew.



Societal Perspective

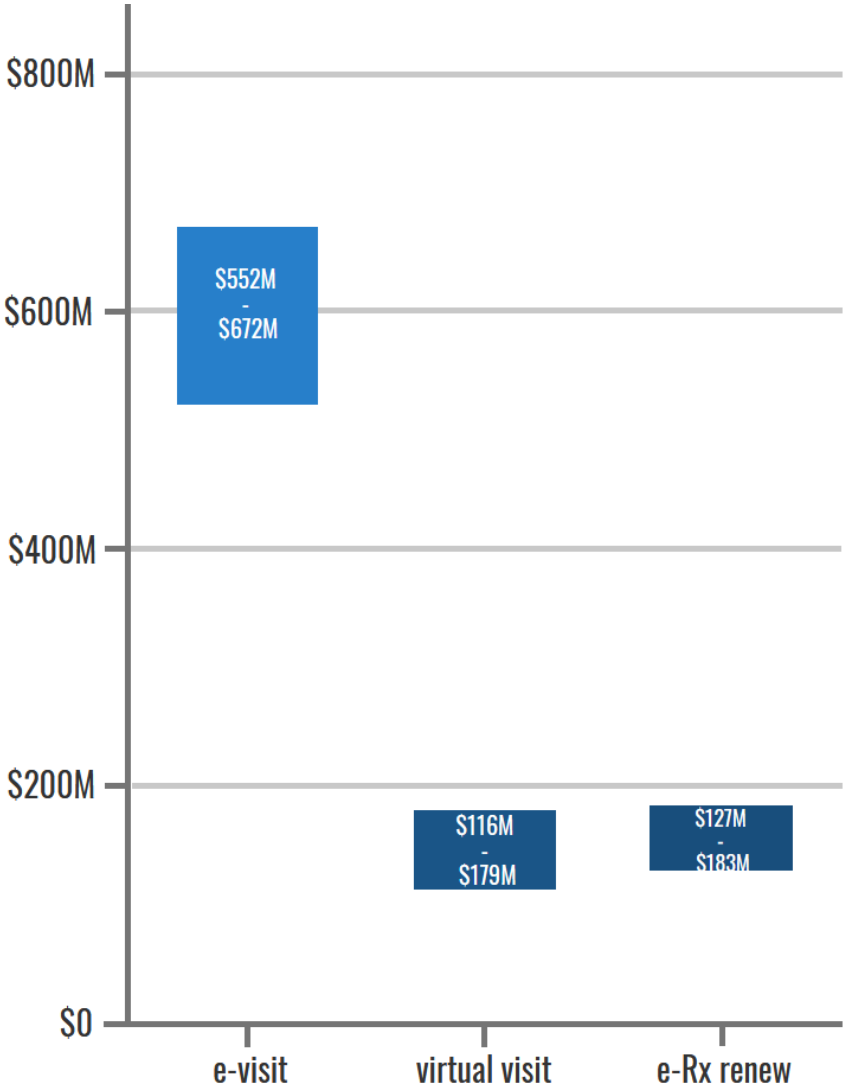
Most evidence regarding PHR use and health outcomes was generated in hospital or integrated care settings in the United States; 1 study was based in Canada. All of the evidence used as inputs into our population health model related to outcomes of interventions targeted specifically to people with diabetes or people with severe and persistent mental illness. Specific outcomes with sufficient evidence included improved health status through greater glycemic control, increased household income through improved life satisfaction, and increased positive health behaviors through improved medication adherence [21,26,40].

The monetary value of those improved health outcomes breaks down as follows in Figure 7, with the majority of the economic benefit realized being driven by the e-Visit functionality.

Multimedia Appendices 9,10, and 11 provide summaries of each of the 3 quantitative models (patient/caregiver, health system, and population health).

XSL•FO
RenderX

Figure 7. Value of improved health outcomes due to personal health record use, by functionality. e-view: electronic view; e-visit: electronic visit; e-Rx renew: electronic prescription renew.



Discussion

Understanding how implementation of PHRs could benefit patients and caregivers, as well as health care systems in high-resource settings such as Canada, is an important step in exploring ways in which digital health technologies can facilitate improved patient engagement and, ultimately, population health outcomes and system transformation strategies for sustainability. We found that Canadians having Web access to their health care information, as well as communication and consultation services with providers via PHR functionalities, generates a substantial economic benefit from patient, health system, and broader societal perspectives. The hierarchy of classification factors we developed presents a practical approach to estimating economic benefit within the field of digital health and patient empowerment research, where questions of who accrues benefits and from which PHR functionalities are likely to inform decisions as to where and how increased access to PHR functionalities should be supported.

Current Benefits

The majority of the current benefits seen in this study to Canadian patients and caregivers were realized in primary care and mental health settings and were resulting from the use of e-Rx renew functionalities . For health systems, the majority of the benefits were realized by citizens' use of e-View functionalities within primary care settings. Comparing these 2 perspectives, the range of estimates for the health systems perspective was much wider than that of the patient/caregiver perspective (1.7%-72% vs 12%-34%, respectively). The relatively large share of benefits driven by e-Views for health systems, particularly in primary care settings, is likely a function of the limited available evidence, as well as the relatively broad population base that reports having access to primary care services in Canada. Patient access to health care information on the Web is therefore clearly linked to cost savings for both patients and health systems.

Evidence for benefits related to improved health outcomes was found in relation to PHR functionalities that enabled greater interaction and communication with health care providers and services: e-Visit, virtual visit, and e-Rx renew. This suggests that having access to technology that allows for more timely interactions between patients and health care providers and clinic administration has added benefits and value for the health of participating patients.

XSL•FO
RenderX

Potential Benefits

For patient and health system perspectives, potential benefits were estimated using 2 different methods. The first method projected increases in adoption rates, and the second method relied on evidence from health management organizations in the United States that had implemented PHRs with integrated communication and clinical consultation e-service functionalities (e-Visit and virtual visits) [42,43]. Although such integrated PHR models are not yet available in Canada, the health outcomes reported resulted from patient use of PHR consultation and communication functionalities and may indeed be applicable to health outcomes in the Canadian context and that of other international health systems. Evidence from current and emerging models of care provides insight into key factors that facilitate health systems modernization to advance clinic/clinician adoption and integration of virtual care electronic services (e-Services), such as remuneration of health care providers.

The majority of potential improvement in health outcomes in our model is driven by increased integration of e-Services within health care organizations and by increased adoption of PHRs by older adults, particularly those in long-term care facilities. We were able to establish a direct link between increased life satisfaction for adults with severe and persistent mental illness who used PHRs compared with those who did not, with an equivalent increase in household income. This contributes to the literature by creating an empirical association between PHR use, patient activation and engagement [18], well-being, and ultimately economic benefit to populations accessing health care services.

Process Matters—Effectiveness and Value of Personal Health Records in Context

Our findings are in line with other studies measuring the value of PHR use compared with business as usual or the absence of a PHR initiative within a health care organization [22,23,44]. There is added value of PHR implementation and adoption by patients and health care organizations, compared with settings with a lack of PHR access, but the degree to which realization of economic value offsets costs of investment and implementation of PHRs is difficult to assess. Similarly, it is challenging to understand the opportunity cost of investing in PHR implementation when benefit estimation perspectives and those sectors investing resources may not be wholly connected.

Critical Mechanisms and Factors of Success

Ultimately, it is likely that PHRs with integrated communication and consultation functionalities will become business as usual in many health care settings and the suite of e-Services available will broaden to include a variety of automated machine learning and artificial intelligence features to support and advance health and wellness needs and facilitated interactions with care providers, organizations, and resources. As availability and adoption become more widespread, factors informing successful implementation, such as integration within health system workflows and provider remuneration structures, will become increasingly important.

Limitations

We drew on evidence synthesized from the peer-reviewed literature as well as benefits evaluations of PHR implementation sites in Canada to develop 3 models estimating benefits of PHR use to Canadians and Canadian health systems. As such, our models are limited by the quality of the evidence generated from these studies. Variability in the number of respondents comprising study sample sizes, as well as the typical absence of a comparison group or counterfactual, also limits the generalizability of our estimates and may result in the over- or underestimation of benefits across the 3 models.

To estimate the value of PHR use from the patient/caregiver perspectives, we valued indirect costs using median income estimates in Canada and 50% of median hourly income to calculate nonlabor costs. It is likely that the distribution of income varies greatly across participants of the studies from which model inputs were drawn. As those who use PHRs tend to have higher income, this likely underestimates the value to current PHR users; however, it may more realistically represent the broader population who could potentially adopt PHRs.

To obtain a Canada-wide set of estimates, we used average costs for health services in Canada; however, these were weighted by the total Canadian population and appropriate adoption rates.

Finally, we did not include the cost of investment in PHR development, implementation, and evaluation, therefore, we do not make any claims about the relative benefit with respect to overall cost. We do, wherever possible, account for variable costs by using only the marginal benefit experienced by health care organizations to calculate estimates of economic value by perspective.

Implications for Research and Practice

PHRs with high levels of direct interaction with care teams showed promising potential value if broadly implemented in the Canadian context with a particular focus on advancing adoption among certain clinical populations—namely, high-frequency health system users and their caregivers. Gaps in evidence include information about PHR use and related outcomes across PHR functions, care settings, and patient populations (see Table 3). From the patient/caregiver perspective, the care settings of studies that generated value-based estimates across e-Services did not include in-patient hospital care. Further research is needed to understand the value to patients and caregivers of PHR use when interacting with acute in-patient care settings. From the health system perspective, gaps in evidence available for estimating values differed according to PHR functionality. To gain a more fulsome understanding of how viewing medical records benefits both patients and health systems, research is needed to evaluate how this may differ across primary care versus specialized in-patient and outpatient hospital settings. There is also a paucity of information available concerning e-Visit, virtual visit, and e-Rx renew service use and related benefits. Finally, an important area for further research is understanding if and how PHR functionalities—particularly virtual visit—influence health system utilization and health outcomes for individuals with chronic conditions, especially older cohorts of adults.

XSL•FO
RenderX

Table 3. Current gaps in evidence related to personal health records.

Personal health record functionality	Electronic visit	Virtual visit	Electronic prescription renew
Care settings with evidence included	Community-based mental health services	Primary care	Community-based mental health and outpatient hospital services
Lack of evidence to inform model	Primary and outpatient specialist care	Outpatient specialist care	Primary care
Priority research area	Yes	Yes	Yes

Overall, more evidence is needed to expand the scope of benefits to include care settings outside of primary care and specialized mental health services, as well as settings advancing access to connected care e-service functionalities (beyond e-View). Understanding the different ways in which different populations benefit—and in which care settings—from saving resources because of PHR use will require further data development, especially regarding:

1. How caregivers use PHRs and how they benefit,
2. How geographic location and proximity to health services influences value for patients,
3. How key determinants of technology adoption may influence who and how patients, caregivers, and health systems use PHRs, and how this relates to any outcomes experienced, and
4. How different populations with high potential value approach PHR use and technology use more broadly.

Conclusions

There is clear value to patients, health systems, and society of patients having Web-based access to their health care information and to consultation and communication e-Services with their providers and care organizations. Increasing this value and benefit would include bridging the digital divide and helping to facilitate training for both patients/caregivers and providers/organizations. To our knowledge, this is the first study synthesizing and estimating the value-based outcomes of PHR adoption and use across multiple sites at a national level. In addition, we feel we have contributed to the field by generating care setting and patient population–specific estimates across a range of adoption rates, effectiveness, and cost levels to explore how targeting various populations (eg, high-volume users, adults with severe and persistent mental illness, older adults in long-term care facilities, and individuals with chronic conditions) could yield further economic and health benefits to patients, caregivers, and health systems. Finally, our findings point to priority areas for new research that can allow for robust economic evaluations of PHRs, guide strategic health system policy and strategy, and identify empirical links between PHR use and cost savings, as well as determinants and mediators of PHR use and individual/population health outcomes across modernizing health care systems around the world.

Acknowledgments

Canada Health Infoway Inc funded this study. Canada Health Infoway Inc also covered the publication costs for this study. Canada Health Infoway Inc is an independent, not-for-profit organization funded by the federal government that works to accelerate the development, adoption, and effective use of digital health across Canada. The funders had no role in the extraction of data from source evidence or specification of model estimates. The funders did review and validate model estimates. CL (affiliated with Canada Health Infoway) was a member of the research team and was involved in study design, interpretation, and manuscript development.

Conflicts of Interest

CL is a Director at Canada Health Infoway, a federally funded national organization advancing citizen's access to their health information and e-Services across Canada.

Multimedia Appendix 1

Source list, patient/caregiver perspective.

[XLSX File (Microsoft Excel File), 14KB - jmir_v21i6e12277_app1.xlsx]

Multimedia Appendix 2

Source list, health system perspective.

[XLSX File (Microsoft Excel File), 14KB - jmir_v21i6e12277_app2.xlsx]

Multimedia Appendix 3

Studies and outcomes included for patient/caregiver perspective.

XSL•FO
RenderX

[DOCX File, 18KB - jmir_v21i6e12277_app3.docx]

Multimedia Appendix 4

Studies and outcomes included in health system perspective.

[DOCX File, 18KB - jmir_v21i6e12277_app4.docx]

Multimedia Appendix 5

Studies and outcomes included in population health perspective.

[DOCX File, 17KB - jmir_v21i6e12277_app5.docx]

Multimedia Appendix 6

Narrative summary of patient/caregiver perspective estimates.

[DOCX File, 57KB - jmir_v21i6e12277_app6.docx]

Multimedia Appendix 7

Narrative summary of health system perspective estimates.

[DOCX File, 31KB - jmir_v21i6e12277_app7.docx]

Multimedia Appendix 8

Narrative summary of population health perspective estimates.

[DOCX File, 31KB - jmir_v21i6e12277_app8.docx]

Multimedia Appendix 9

Quantitative model summary, patient/caregiver.

[XLSX File (Microsoft Excel File), 25KB - jmir_v21i6e12277_app9.xlsx]

Multimedia Appendix 10

Quantitative model summary, health system.

[XLSX File (Microsoft Excel File), 14KB - jmir_v21i6e12277_app10.xlsx]

Multimedia Appendix 11

Quantitative model summary, population health.

[XLSX File (Microsoft Excel File), 9KB - jmir_v21i6e12277_app11.xlsx]

References

1. Beun J. Electronic healthcare record; a way to empower the patient. Int J Med Inform 2003 Mar;69(2-3):191-196. [Medline: 12810123]
2. Juciute R. ICT implementation in the health-care sector: effective stakeholders' engagement as the main precondition of change sustainability. AI & Soc 2007 Sep 25;23(1):131-137. [doi: 10.1007/s00146-007-0168-4]
3. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan 18;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
4. Pearce C, Bainbridge M. A personally controlled electronic health record for Australia. J Am Med Inform Assoc 2014;21(4):707-713 [FREE Full text] [doi: 10.1136/amiajnl-2013-002068] [Medline: 24650635]
5. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Dec 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
6. Schneider H, Hill S, Blandford A. Patients know best: qualitative study on how families use patient-controlled personal health records. J Med Internet Res 2016 Feb 24;18(2):e43 [FREE Full text] [doi: 10.2196/jmir.4652] [Medline: 26912201]
7. Canada Health Infoway. Connecting Patients for Better Health. Montreal, Quebec: Canada Health Infoway; 2016.

XSL•FO
RenderX

8. Earnscliffe SG, Canada HI. 'What Canadians Think': Canadians' Perspectives on Privacy of Personal Health Information in the Context of Digital Health. Montreal, Quebec: Canada Health Infoway; 2017.
9. Nykänen P, Brender J, Talmon J, de Keizer N, Rigby M, Beuscart-Zephir M, et al. Guideline for good evaluation practice in health informatics (GEP-HI). Int J Med Inform 2011 Dec;80(12):815-827. [doi: 10.1016/j.ijmedinf.2011.08.004] [Medline: 21920809]
10. Rigby M, Georgiou A, Hyppönen H, Ammenwerth E, de Keizer N, Magrabi F, et al. Patient portals as a means of information and communication technology support to patient-centric care coordination–the missing evidence and the challenges of evaluation. A joint contribution of IMIA WG EVAL and EFMI WG EVAL. Yearb Med Inform 2015 Aug 13;10(1):148-159 [FREE Full text] [doi: 10.15265/IY-2015-007] [Medline: 26123909]
11. Bassi J, Lau F. Measuring value for money: a scoping review on economic evaluation of health information systems. J Am Med Inform Assoc 2013;20(4):792-801 [FREE Full text] [doi: 10.1136/amiajnl-2012-001422] [Medline: 23416247]
12. National Alliance for Health Information Technology. HITECH Answers. 2008 Apr. Defining Key Health Information Technology Terms URL: http://www.hitechanswers.net/wp-content/uploads/2013/05/NAHIT-Definitions2008.pdf [accessed 2019-01-30] [WebCite Cache ID 75ooudC6h]
13. Kaelber D, Shah S, Vincent A, Pan E, Hook J, Johnston D, et al. The Value of Personal Health Records. Charlestown: Acta Musei; 2008.
14. Werder M. Health information technology: a key ingredient of the patient experience. Patient Exp J 2015;2(1):143-147 [FREE Full text]
15. Carayon P, Hoonakker P, Cartmill R, Hassol A. Using Health IT in Practice Redeisgn: Impact of Health IT on Workflow. Rockville, MD: Agency for Healthcare Research and Quality; 2015.
16. Emani S, Yamin CK, Peters E, Karson AS, Lipsitz SR, Wald JS, et al. Patient perceptions of a personal health record: a test of the diffusion of innovation model. J Med Internet Res 2012 Nov 5;14(6):e150 [FREE Full text] [doi: 10.2196/jmir.2278] [Medline: 23128775]
17. Kaelber D, Jha A, Johnston D, Middleton B, Bates D. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008;15(6):729-736 [FREE Full text] [doi: 10.1197/jamia.M2547] [Medline: 18756002]
18. Risling T, Martinez J, Young J, Thorp-Froslie N. Evaluating patient empowerment in association with eHealth technology: scoping review. J Med Internet Res 2017 Dec 29;19(9):e329 [FREE Full text] [doi: 10.2196/jmir.7809] [Medline: 28963090]
19. Institute for Healthcare Improvement. 2009. Triple Aim - Concept Design URL: http://www.ihi.org/Engage/Initiatives/TripleAim/Documents/ConceptDesign.pdf [accessed 2018-09-20] [WebCite Cache ID 72Zma9lKI]
20. Kontos E, Blake K, Chou W, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
21. Ministry of Health and Long Term Care. Patients First: A Proposal to Strengthen Patient-Centred Health Care in Ontario. Toronto, ON: Ministry of Health and Long Term Care; 2015.
22. Finkelstein J, Knight A, Marinopoulos S, Gibbons M, Berger Z, Aboumatar H, et al. Enabling patient-centered care through health information technology. Evid Rep Technol Assess (Full Rep) 2012 Jun(206):1-1531. [Medline: 24422882]
23. Vannieuwenborg F, van der Auwermeulen T, van Ooteghem J, Jacobs A, Verbrugge S, Colle D. Evaluating the economic impact of smart care platforms: qualitative and quantitative results of a case study. JMIR Med Inform 2016 Oct 31;4(4):e33 [FREE Full text] [doi: 10.2196/medinform.5012] [Medline: 27799137]
24. Amante DJ, Hogan TP, Pagoto SL, English TM. A systematic review of electronic portal usage among patients with diabetes. Diabetes Technol Ther 2014 Nov;16(11):784-793. [doi: 10.1089/dia.2014.0078] [Medline: 24999599]
25. Chung K, Cheng S, Leonard KJ, Dalziel S. Access to test results. ElectronicHealthcare 2012;11(2):17-23 [FREE Full text]
26. Lyles CR, Sarkar U, Schillinger D, Ralston JD, Allen JY, Nguyen R, et al. Refilling medications through an online patient portal: consistent improvements in adherence across racial/ethnic groups. J Am Med Inform Assoc 2016 Apr;23(e1):e28-e33 [FREE Full text] [doi: 10.1093/jamia/ocv126] [Medline: 26335983]
27. Hibbard J. Patient Activation: Improving Health Outcomes and Reducing Costs. Eugene, Oregon: University of Oregon; 2015.
28. Hibbard JH, Greene J, Becker ER, Roblin D, Painter MW, Perez DJ, et al. Racial/ethnic disparities and consumer activation in health. Health Aff (Millwood) 2008;27(5):1442-1453. [doi: 10.1377/hlthaff.27.5.1442] [Medline: 18780935]
29. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood) 2015 Mar;34(3):431-437. [doi: 10.1377/hlthaff.2014.0452] [Medline: 25732493]
30. Sorondo B, Allen A, Fathima S, Bayleran J, Sabbagh I. Patient portal as a tool for enhancing patient experience and improving quality of care in primary care practices. EGEMS (Wash DC) 2016;4(1):1262 [FREE Full text] [doi: 10.13063/2327-9214.1262] [Medline: 28203611]
31. Sarkar U, Lyles CR, Parker MM, Allen J, Nguyen R, Moffet HH, et al. Use of the refill function through an online patient portal is associated with improved adherence to statins in an integrated health system. Med Care 2014 Mar;52(3):194-201 [FREE Full text] [doi: 10.1097/MLR.0000000000000069] [Medline: 24374412]

XSL•FO
RenderX

32. Reiners F, Sturm J, Bouw LJ, Wouters EJ. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Dec 21;16(4):e645 [FREE Full text] [doi: 10.3390/ijerph16040645] [Medline: 30795623]
33. van Deursen AJ, van Dijk JA. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. New Media Soc 2019 Feb;21(2):354-375 [FREE Full text] [doi: 10.1177/1461444818797082] [Medline: 30886536]
34. Sadasivaiah S, Lyles CR, Kiyoi S, Wong P, Ratanawongsa N. Disparities in patient-reported interest in web-based patient portals: survey at an urban academic safety-net hospital. J Med Internet Res 2019 Mar 26;21(3):e11421 [FREE Full text] [doi: 10.2196/11421] [Medline: 30912747]
35. Antonio MG, Petrovskaya O. Towards devloping an eHealth equity conceptual framework. In: Lau F, Bartle-clar J, Bliss G, Borycki E, Courtney K, Kuo A, et al, editors. Improving Usability, Safety And Patient Outcomes With Health Information Technology: From Research To Practice. Amsterdam: IOS Press; 2019.
36. O'Reilly D, Holbrook A, Blackhouse G, Troyan S, Goeree R. Cost-effectiveness of a shared computerized decision support system for diabetes linked to electronic medical records. J Am Med Inform Assoc 2012;19(3):341-345 [FREE Full text] [doi: 10.1136/amiajnl-2011-000371] [Medline: 22052900]
37. DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. J Manage Inf Syst 2003;19(4):9-30. [doi: 10.1080/07421222.2003.11045748]
38. Canada Health Infoway. Canada Health Infoway. Montreal, Quebec: Canada Health Infoway; 2012 Apr. Canada Health Infoway Benefits Evaluation Indicators Technical Report URL: https://www.infoway-inforoute.ca/en/component/edocman/450-benefits-evaluation-indicators-technical-report-version-2-0/view-document?Itemid=101 [accessed 2018-09-20] [WebCite Cache ID 72Zniiumn]
39. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes. Third Edition. Oxford: Oxford University Press; 2005.
40. Helliwell JF, Huang H. How's the job? Well-being and social capital in the workplace. Ind Labour Relations Rev 2010 Jan;63(2):205-227. [doi: 10.1177/001979391006300202]
41. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan 27;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
42. Byrne CM, Mercincavage LM, Bouhaddou O, Bennett JR, Pan EC, Botts NE, et al. The Department of Veterans Affairs' (VA) implementation of the Virtual Lifetime Electronic Record (VLER): findings and lessons learned from Health Information Exchange at 12 sites. Int J Med Inform 2014 Aug;83(8):537-547. [doi: 10.1016/j.ijmedinf.2014.04.005] [Medline: 24845146]
43. King G, Maxwell J, Karmali A, Hagens S, Pinto M, Williams L, et al. Connecting families to their health record and care team: the use, utility, and impact of a client/family health portal at a children's rehabilitation hospital. J Med Internet Res 2017 Apr 6;19(4):e97 [FREE Full text] [doi: 10.2196/jmir.6811] [Medline: 28385680]
44. Emont S. Measuring the Impact of Patient Portals?: What the Literature Tells Us. Oakland: California HealthCare Foundation; May 2011.

Abbreviations

EHR: electronic health record
e-Rx renew: electronic prescription renew
e-Services: electronic services
e-View: electronic view
e-Visit: electronic visit
PHR: personal health record

Edited by G Eysenbach; submitted 20.09.18; peer-reviewed by T Risling, J Maxwell, C Byrne; comments to author 20.12.18; revised version received 12.02.19; accepted 02.04.19; published 06.06.19.

Please cite as:
Hackett C, Brennan K, Smith Fowler H, Leaver C
Valuing Citizen Access to Digital Health Services: Applied Value-Based Outcomes in the Canadian Context and Tools for Modernizing Health Systems
J Med Internet Res 2019;21(6):e12277
URL: https://www.jmir.org/2019/6/e12277/
doi:10.2196/12277
PMID:31172965

XSL•FO
RenderX

©Christina Hackett, Kelsey Brennan, Heather Smith Fowler, Chad Leaver. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Internet of Things Buttons to Measure and Respond to Restroom Cleanliness in a Hospital Setting: Descriptive Study

Peter R Chai[1,2,3], MA, MD; Haipeng Zhang[3,4,5], DO; Guruprasad D Jambaulikar[1], MBBS; Edward W Boyer[1,2,3], MD, PhD; Labina Shrestha[6], MA; Loay Kitmitto[6], BA; Paige G Wickner[3,7], MD; Hojjat Salmasian[3,7], PhD; Adam B Landman[1,3,4], MD

[1]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States

[2]The Fenway Institute, Boston, MA, United States

[3]Harvard Medical School, Boston, MA, United States

[4]Digital Innovation Hub, Brigham and Women's Hospital, Boston, MA, United States

[5]Department of Psychosocial Oncology and Palliative Care, Dana Farber Cancer Institute, Boston, MA, United States

[6]Environmental Services, Brigham and Women's Hospital, Boston, MA, United States

[7]Department of Quality and Safety, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
Peter R Chai, MA, MD
Department of Emergency Medicine
Brigham and Women's Hospital
75 Francis St
Boston, MA, 02411
United States
Phone: 1 617 732 5640
Email: pchai@bwh.harvard.edu

Abstract

Background: Restroom cleanliness is an important factor in hospital quality. Due to its dynamic process, it can be difficult to detect the presence of dirty restrooms that need to be cleaned. Using an Internet of Things (IoT) button can permit users to designate restrooms that need cleaning and in turn, allow prompt response from housekeeping to maintain real-time restroom cleanliness.

Objective: This study aimed to describe the deployment of an IoT button–based notification system to measure hospital restroom cleanliness reporting system usage and qualitative feedback from housekeeping staff on IoT button use.

Methods: We deployed IoT buttons in 16 hospital restrooms. Over an 8-month period, housekeeping staff received real-time notifications and responded to button presses for restroom cleaning. All button presses were recorded. We reported average button usage by hospital area, time of day, and day of week. We also conducted interviews with housekeeping supervisors and staff to understand their acceptance of and experience with the system.

Results: Over 8 months, 1920 requests to clean restrooms in the main hospital lobby and satellite buildings were received. The hospital lobby IoT buttons received over half (N=1055, 55%) of requests for cleaning. Most requests occurred in afternoon hours from 3 PM to midnight. Requests for cleaning remained stable throughout the work week with fewer requests occurring over weekends. IoT button use was sustained throughout the study period. Interviews with housekeeping supervisors and staff demonstrated acceptance of the IoT buttons; actual use was centered around asynchronous communication between supervisors and staff in response to requests to clean restrooms.

Conclusions: An IoT button system is a feasible method to generate on-demand request for restroom cleaning that is easy to deploy and that users will consistently engage with. Data from this system have the potential to enable responsive scheduling for restroom service and anticipate periods of high restroom utilization in a hospital.

(J Med Internet Res 2019;21(6):e13588) doi:10.2196/13588

KEYWORDS

operations research; wireless technology; hygiene; toilet facilities; workflow

XSL•FO
RenderX

Introduction

Background

Cleanliness of restrooms frequently serves as a gauge of an organization's ability to maintain cleanliness throughout all of its facilities. The maintenance of restroom cleanliness is important for 3 reasons. First, dirty restrooms are frequently cited as a source of complaints in many industries. In a 2010 poll, individuals identified dirty restrooms as the top reason to avoid a restaurant with the most common complaints being about clogged toilets, foul odors, out-of-stock supplies, and broken soap or paper dispensers [1,2]. Hospitals are also frequently cited as a source of dirty restrooms; a 2004 survey of 86,000 patients in the United Kingdom showed that only 48% of individuals thought their hospital restrooms were clean [3]. Second, the cleanliness of hospital restrooms is associated with the quality of hospital. In national focus groups designed to inform the Hospital Consumer Assessments of Healthcare Providers and Systems Hospital Survey, 15 of 16 focus groups identified cleanliness of restrooms as a gauge through which participants would measure hospital quality [4]. Similarly, focus groups in Greece demonstrated that patients use cleanliness of restrooms as an indicator of overall quality of the hospital [5]. Third, in health care facilities where nosocomial infections continue to be of a great concern, dirty restrooms can be a source of bacterial pathogens from biowaste. Toilets left unclean may lead to aerosolization of biowaste after flushing, serving as a potential source of infection, especially in immunocompromized individuals [6-8].

Maintaining Restroom Cleanliness in the Hospital Setting

Hospitals utilize various systems to maintain restroom cleanliness from scheduled, periodic cleaning of high-traffic restrooms to on-demand cleaning based on user feedback. These strategies are symbiotic—ideally, a restroom is cleaned on a schedule based on the number of user requests. Restroom users may request restroom cleaning by notifying nearby staff or calling a posted phone extension to report a dirty restroom. Novel digital techniques to deliver just-in-time notification of restroom cleanliness may use smartphones to report dirty restrooms via a quick response code (QR code), short message service (SMS) short code, or a dedicated app. Digital solutions have the advantage of automating the process of data collection into a central database that can be queried and analyzed to anticipate staffing needs or times when restrooms are most dirty. Yet, existing digital solutions still require user input (eg, using an individual's smartphone to access an app or to scan a QR code), or require programming of an app, which may be barriers to adoption of such technologies.

An Internet of Things (IoT) button may serve as a simpler, acceptable method to deliver just-in-time notifications of dirty restrooms [9]. An IoT button is a small, electronic device that can be programmed to deliver a customized message to the end user via a wireless network when pressed. In this study, we describe the feasibility and acceptability of deploying an IoT button-based notification system to measure restroom cleanliness in our hospital. We also report descriptive statistics on the use of IoT buttons and share qualitative feedback from housekeeping supervisors and staff on use of IoT buttons.

Methods

Overview

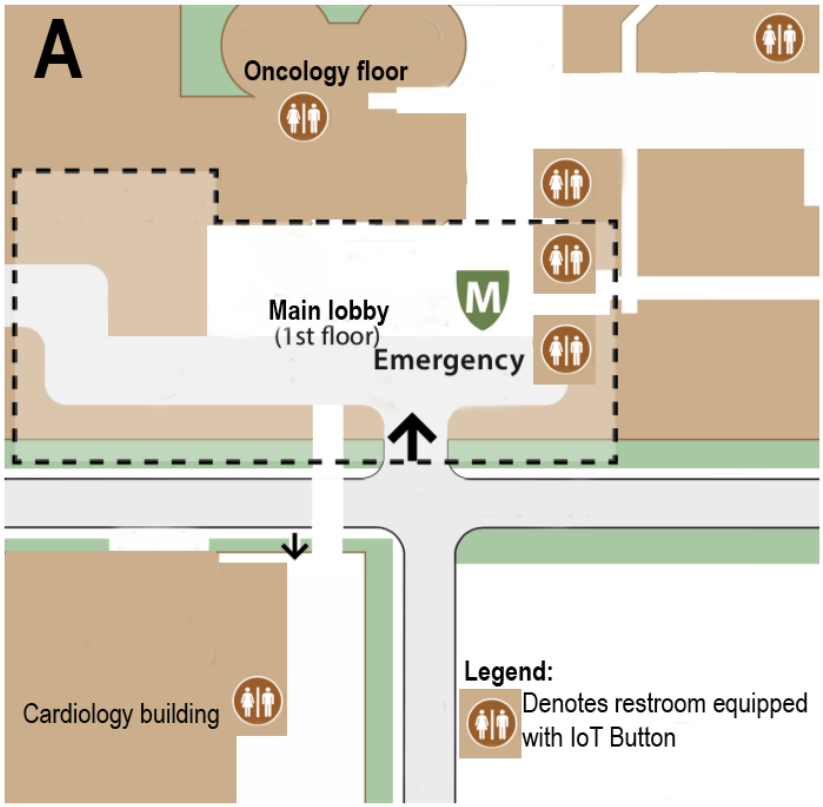
We utilized a modified IoT button (GoButton, Visybl Inc, Germantown, MD, USA) configured to deliver just-in-time messages to the housekeeping staff regarding restroom cleaning requests and grounded our investigation in a plan-do-study-act model [10,11]. We conducted our project at a large, urban, academic quaternary care hospital with 763 beds. Our hospital receives approximately 30,000 visitors daily and emergency department (ED) has an annual volume of 60,000 patients.

Before the implementation of the IoT buttons, our housekeeping staff used to clean designated restrooms on a schedule that reflected estimated daily restroom use. For example, our highest volume restrooms in the main lobby have a dedicated housekeeper who inspects the restroom on an hourly basis, whereas an inpatient floor restroom is cleaned once daily.

We initially selected 9 high-traffic public restrooms throughout the hospital in collaboration with housekeeping leadership (Figure 1). Selection criteria were based on the level of foot traffic and historical high utilization of housekeeping services resources in these restrooms. We also selected the public restrooms on the oncology floor as restroom cleanliness on this floor had previously been a priority for housekeeping staff due to the presence of immunocompromized patients. After 1 month at the request of ED and housekeeping leadership, we also equipped each restroom in the ED (a total of 7) with the IoT button because the cleanliness of these restrooms was identified as an important contributor to patient satisfaction in our hospital. Overall, we included 16 restrooms in our data analysis.

XSL•FO
RenderX

Figure 1. Floor map of major sections of the main hospital building with designated restrooms where Internet of Things buttons were displayed.



Internet of Things Button Programming and Deployment

We mapped each IoT button to a single restroom using the vendor's Web-based administrator interface [9]. We also used this system to program the IoT button to deliver notifications via SMS messages to housekeeping supervisors based on specific button actions (Figure 2). We designated a single press as a notification that restrooms needed cleaning. In addition, a double press (2 presses of the IoT button within 1s of each other) delivered a notification that the restroom was cleaned by housekeeping staff. Each action—single or double press—was automatically logged on a cloud-hosted, shared spreadsheet. As our housekeeping standard is to respond and clean a restroom within 30 min of a request, we created a button lockout period of 30 min where only the first single press during a 30-min period would deliver a message to housekeeping supervisors. Single presses during the 30-min lockout period were still logged by the interface.

We conducted periodic in-service training for housekeeping staff using a demonstration IoT button. Housekeeping supervisors introduced the device and its intended use to housekeeping staff at daily morning staff meetings. We also conducted in-service training in the evening for housekeepers working the overnight shift. After demonstration of the IoT button, we allowed housekeeping staff to inspect and use the demonstration button. The demonstration button was kept in the housekeeping supervisor's office so that housekeeping staff could readily access and retrain as needed.

IoT buttons were mounted in readily accessible areas on restroom walls with public facing signage asking users to press the IoT button once if the restroom needed service (Figure 3). Each button press delivers an SMS message to the designated housekeeping supervisor responsible for restrooms in a specific region who would then dispatch housekeeping staff to assess and clean the restroom. Housekeeping staff were trained by the supervisors to double press the IoT button each time they cleaned the restroom in routine cleanings as well in response to button requests. Periodic retraining occurred during onboarding of new housekeeping staff and during daily housekeeping meetings.

At the end of the study period, we conducted face-to-face group interviews grounded in the technology acceptance model with 4 housekeeping supervisors and 2 staff regarding their experience using the IoT button system [12]. Housekeeping staff and supervisors were recruited during regular operational staff meetings. We utilized a standard interview guide aimed at the intended use of the IoT button, experience integrating the IoT button into the existing restroom cleaning workflow, and optimization of the IoT button for future use. The deployment of the IoT buttons and analysis of the log data was a quality improvement initiative, not human subjects' research. For the qualitative interview component of the investigation, we obtained institutional review board exemption.

XSL•FO
RenderX

Figure 2. Process diagram of Internet of Things button activation and response.

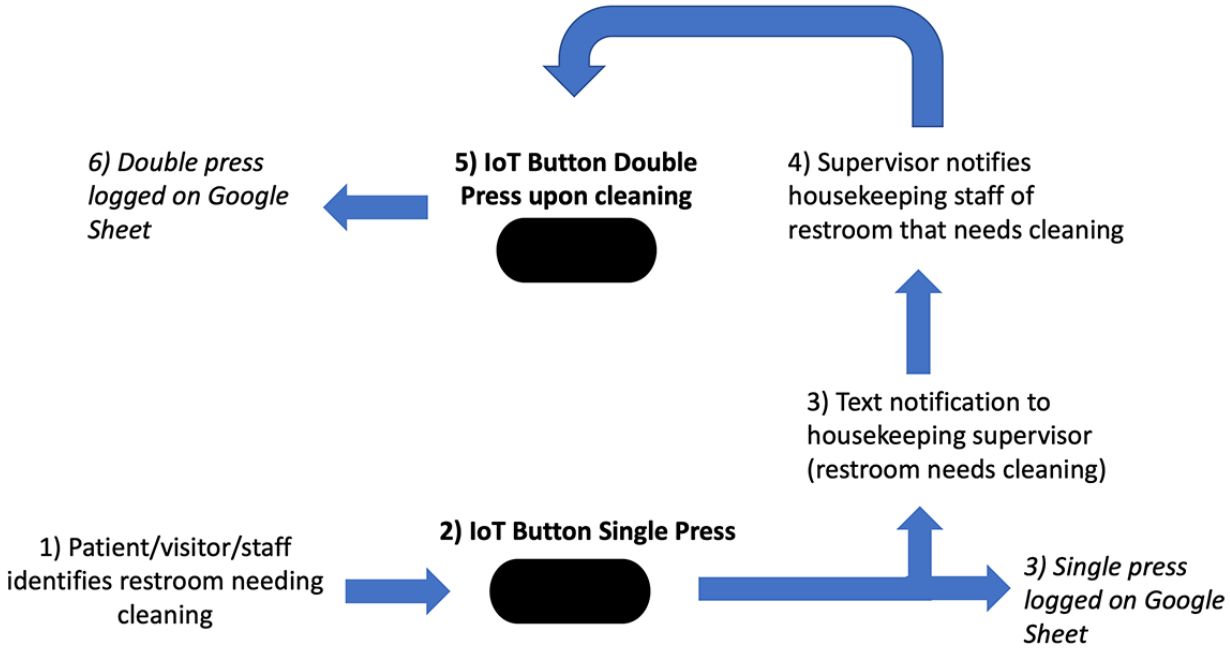


Figure 3. Signage around Internet of Things buttons instructing the public to press the button if they think the restroom needs cleaning.

Data Analysis

We analyzed the audit logs of the IoT buttons, which we obtained from the cloud-based, shared spreadsheet. Audit logs contained the coded location of the IoT button, and the date and time as well as type (single or double press) of the button press. We aggregated all single and double button presses by restroom and also region of the hospital corresponding to the regions of each housekeeping supervisor. We excluded button presses that originated from one IoT button designated as a training device for housekeeping staff. Next, we separated button presses according to housekeeping shifts (7 am to 4 pm , 3 pm to 12 midnight, and 11 pm to 8 am). Presses that occurred during the overlapping shift periods (3 pm to 4 pm, 11 pm to 12 midnight, and 7 am to 8 am) were accounted for in the earlier shift. For example, if a button was activated at 3:30 PM, we counted it as an activation during the 7 am to 4 pm shift. We measured trends of button presses across days of the week. We also calculated a time to next request defined as the number of hours between a single press and a subsequent single press of the same IoT button, as a proxy for how long it took for the bathroom to need cleaning again. We analyzed whether the time to next request was different when cleaning occurred after a request (as evident by a double press of the same IoT button in between the 2 single presses). Data were summarized and reported as mean (SD) and comparisons were made using standard

XSL•FO
RenderX

parametric tests (t test and Chi-squared); we used R version 3.6.0 (The R Foundation, Vienna, Austria) to conduct the analyses [13].

Group interviews were conducted by the study staff. Notes regarding participant responses were taken in real time and findings were discussed by members of the study team. Major themes from these discussions were recorded and presented to the group for review. The purpose of qualitative interviews was to gather formative data regarding the early stage deployment of the IoT Button. As a result, we did not complete a formal applied thematic analysis.

Results

Overview

We collected data from 16 restroom IoT buttons from November 2017 to July 2018 (Table 1). Overall, we recorded a total of 2678 presses. We excluded 70 recorded IoT button presses, which were used during housekeeping staff training. A total of 1920 single-press requests for cleaning were recorded, whereas 688 cleaning confirmations (double-press events) were recorded.

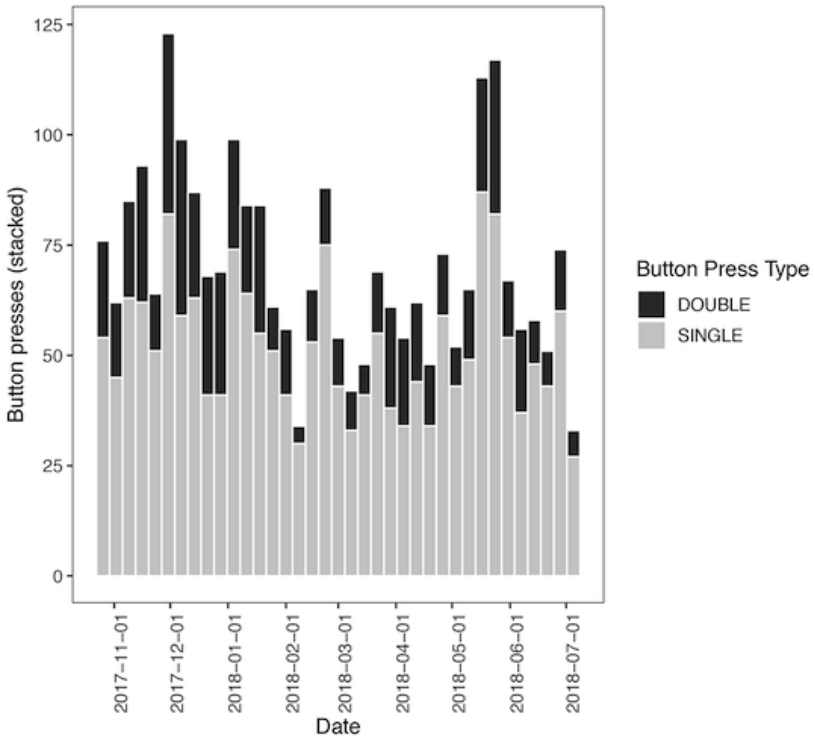
We observed persistent use of IoT buttons during the study period (Figure 4). During the course of the study period, we had to replace 3 IoT buttons in 3 distinct restrooms that were reported lost. We noted on examination that these buttons likely fell off the signage due to lost integrity of the adhesive originally used. This did not impact our analyses, as we reused the previous button identifiers.

We recorded 1055 total requests for cleaning in the main hospital (ie, men and women's restrooms in the main entrance lobby of the hospital), accounting for 55% of all requests recorded during the study. The main hospital lobby restrooms were also responsible for the highest number of confirmed cleanings in the main hospital (393/688, 57.1%). On average, 263 button presses were recorded from buttons in the main hospital during the study period (Table 1). There were 384 requests to clean these restrooms in our cardiology building connected to the main hospital over the course of the study period, accounting for 20% of total requests, and a total of 86 confirmed cleaning episodes (86/688, 12.5%). The oncology floor comprised the fewest requests for cleaning in our sample with only 132 (7%) requests during the study period and 22 confirmed cleaning episodes (3%).

Table 1. Internet of Things button utilization by hospital location over the study period (8 months).

Location	Number of buttons deployed	Average single presses per button (SD)	Average double presses per button (SD)
Main hospital	4	263.8 (179.8)	97.3 (51.0)
Cardiology building	4	96.0 (43.0)	21.5 (5.4)
Oncology floor	1	132.0 (0)	22.0 (0)
Emergency department	7	49.9 (26.1)	31.1 (14.1)
Overall	16	120.0 (22.8)	45.9 (41.5)

Figure 4. Trends of Internet of Things button utilization stratified by type of press. Each column represents 1 week.



XSL•FO
RenderX

We equipped restrooms in the ED and ED waiting room with IoT buttons 1 month after the start of the study. Despite a late deployment, we recorded an average of 49.9 requests for cleaning in the ED over the study period. There was a total of 349 requests for cleaning in the ED during the study period. Of requests in the ED, 28.7% (100/349) originated from the waiting room. There was a total of 197 confirmed cleans within the ED, 85 of which occurred in the waiting room.

Next, we grouped requests for cleaning by shift (Table 2). An average of 120 requests for cleaning occurred per month during the afternoon shift (3 PM to midnight). There was no variation in requests for cleaning during the work week, although there were fewer requests for cleaning during the weekend, defined as Saturday to Sunday (Table 3).

The median time between 2 single-press events of the same IoT button was 15.2 hours (interquartile range 45.5 hours; Figure 5). This time was longer when a cleaning event happened in between 2 single-press events (difference 57 min; 95% CI 9-138 min; P=.008). We also measured IoT button use over time. Outside of an increase in presses 1 month after deployment with the addition of IoT buttons in the ED, we observed a steady number of button presses per month.

Table 2. Average Internet of Things button utilization per month by shift time.

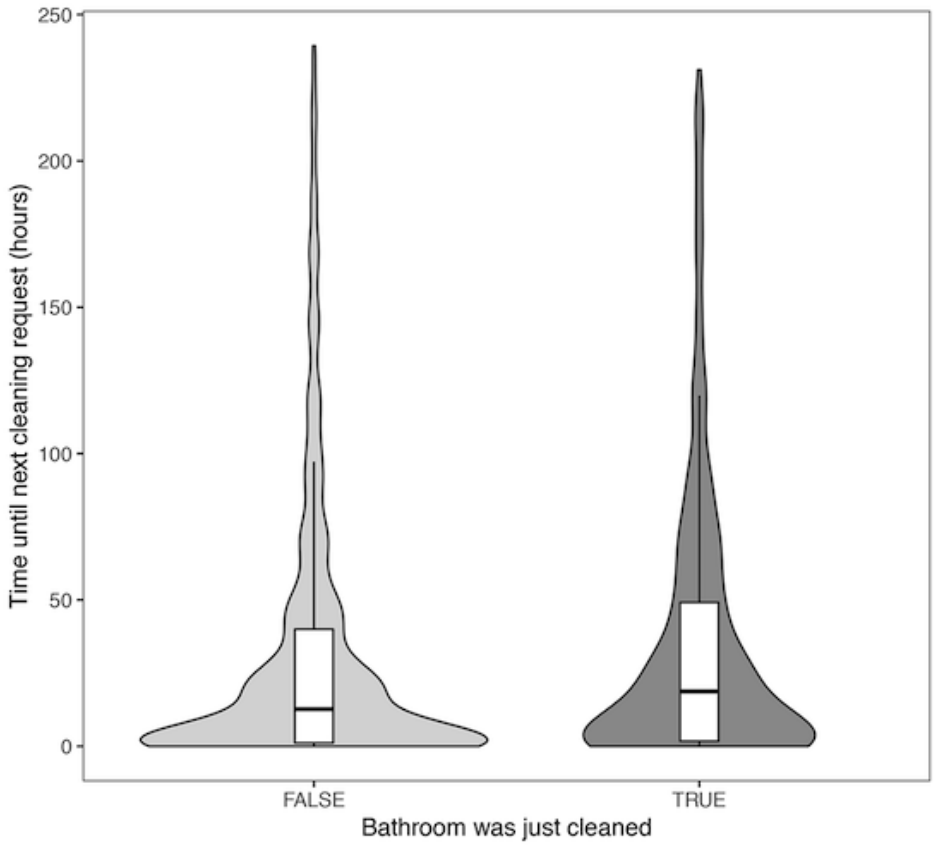
Shift time	Single press, mean (SD)	Double press, mean (SD)
Morning (7 am to 4 pm)	38.9 (18.6)	23.7 (14.9)
Afternoon (3 pm to 12 midnight)	120.0 (53.9)	37.7 (22.9)
Evening (11 pm to 8 am)	33.1 (13.5)	7.4 (5.5)

Table 3. Average Internet of Things button utilization by day of the week.

Day of the week	Single press, mean (SD)	Double press, mean (SD)
Monday	29.8 (14.3)	13.4 (7.0)
Tuesday	37.6 (19.5)	10.7 (4.3)
Wednesday	32.9 (19.7)	10.8 (3.9)
Thursday	32.3 (16.8)	9.5 (6.7)
Friday	30.2 (13.9)	11.4 (6.7)
Saturday	13.0 (6.7)	9.3 (10.0)
Sunday	13.9 (5.3)	9.6 (10.3)

XSL•FO
RenderX

Figure 5. Violin plot demonstrating relationship between an Internet of Things button single press (request for cleaning) and double press (confirmed cleaning request). When confirmed cleaning requests occurred in between two requests for cleaning, the time was longer for a subsequent request. True includes bathrooms with a logged (completed) cleaning event.



Qualitative Response to Internet of Things Button Deployment

We conducted 2 group interviews; one group consisted of 4 housekeeping supervisors and the other group consisted of 2 housekeeping staff. Our environmental services department employs approximately 10 supervisors. Both groups reported high acceptance of the IoT button to gauge the cleanliness of restrooms. Regarding intention to use, supervisors reported that they wanted to use the IoT button to understand when the volume of requests for cleaning was the highest to dispatch housekeeping staff to clean. In addition, supervisors reported that they intended to use the system to understand workflow of their staff. By recording the double presses, supervisors felt like they were being notified that the restroom was cleaned. In actual use, supervisors reported that they recognized the inherent fallacies of the IoT button system—that there was a lack of information as to why the button was being pressed. Supervisors evolved to respond not to individual requests for clean, but requests in aggregate over time. For example, one supervisor reported they received requests for clean on their phone during the day, and if the frequency of requests started to increase, they would call housekeeping staff and direct them to inspect and clean the restroom. Finally, supervisors reported over the course of the study period, they came to rely on the IoT button system to dispatch the housekeeping staff to potential dirty restrooms.

Housekeeping staff were also accepting of the IoT button system. Staff reported that they felt like most requests for cleaning were likely due to public curiosity with the device. When responding to requests for cleaning, staff reported that there was no context to what was dirty in the restroom. This resulted in staff inspecting and cleaning all aspects of a restroom when a supervisor called. Despite this, staff reported that they did not mind being called by supervisors in response to cleaning requests—staff reported that they felt it was their job to clean restrooms, and if they arrived at a restroom that was reported dirty but found to be clean, then their task would be complete. Housekeeping staff also noted that they frequently forgot to double press the button to indicate that they cleaned the restroom because it represented an additional step in their workflow, but if their supervisor had called and dispatched them to clean the restroom, they would double press the IoT button to show supervisors that their job was complete. Overall, housekeeping staff reported acceptance of the IoT button system during the study period and reported that the system was a tool that helped them keep restrooms cleaner.

Discussion

Principal Findings

Our data demonstrate that an IoT button is a feasible method to gauge the public perception of restroom cleanliness in a hospital. Over the course of 1 year, we demonstrated persistent use of the IoT button system. Housekeeping supervisors and staff reported that real-time data from the IoT button were helpful in maintaining the cleanliness of restroom. By the end

XSL•FO
RenderX

of the investigation, both staff and supervisors had integrated the IoT button system into their daily workflow demonstrating its adoptability. In addition, we only had to replace 3 IoT buttons during the study period—these buttons were lost likely due to degradation of adhesive that was used to affix the buttons to placards. These results are important because they demonstrate that a low-cost, IoT button solution can be deployed at a large hospital setting, the public will continue to engage with the system, and meaningful, actionable data can be generated to assist housekeeping with daily services.

The IoT button system allowed us to measure restroom cleaning requests in a dynamic manner. Due to the simplicity of the device, we experienced more requests from individuals using the restrooms instead of confirmation of cleaning events by housekeeping staff; 71.29% (1920/2693) of presses of the button came from requests for a restroom to be cleaned. Restrooms in places with higher traffic such as those in the main hospital received higher requests for cleaning in comparison with restrooms on an inpatient ward. These findings suggest that the IoT button system can enable contextual awareness of restroom cleanliness to enable real-time housekeeping staffing levels based on requests for cleaning [14]. For example, during times of high public traffic, if there is a real-time increase in requests for clean in one part of the hospital, this could be an indication to housekeeping leadership to move a housekeeper to that area from one where there are less requests for clean. In this manner, we think that future iterations of the IoT button system can help guide and even predict how to efficiently deploy housekeeping staff in the hospital.

Of note, the IoT button data indicated that when a restroom was confirmed cleaned by housekeeping staff it would take longer for a subsequent cleaning request to be requested at that restroom. Although the observed difference was only 57 min, it should be noted that many restrooms are already cleaned multiple times a day, and a larger difference in the data may be practically infeasible. In addition, staff reported that if they cleaned restrooms based on their set schedule, they frequently did not log these cleaning episodes on the IoT button system. With increased adherence to logging cleaning episodes on the IoT button system, this system could be used to understand response times to requests for cleaning. More frequent in-service training of staff members may help improve reporting of cleaning episodes. The IoT button system may also help not only respond to dirty restrooms, but over time, maintain cleaner restrooms, and improve perceived quality of a hospital [15]. This analysis is likely limited by the fact that less than 2 cleaning events were recorded using a double press for 96% of location-days. With improved training of staff and more accurate recording of double presses, the IoT button system may be used to maintain cleaner restrooms.

Our formative qualitative data support end-user adoption and integration of the IoT button into the daily workflow of our housekeeping staff, despite its introduction as a quality improvement project. Major identified themes surrounded the ability to access and transmit restroom data in real time, and the use of the button to report cleaning episodes. Importantly, housekeeping staff were willing and able to double press the IoT button as they responded to requests by their supervisors to clean restrooms demonstrating the importance of completing feedback loops and reinforcing their daily task of helping to maintain clean restrooms [16]. Despite identified limitations with the workflow and data, supervisors continued to utilize the IoT button to understand the pulse of restroom cleanliness during the day whereas housekeeping staff found the IoT button an easy method in which to communicate completion of a task (cleaned restroom) to supervisors.

Challenges in Implementing the Internet of Things Button System

We experienced data security issues as we considered scaling the IoT button system. The IoT button version we used only supports Wi-Fi Protected Access 2 (WPA2) encryption, which only requires a single preshared key (PSK) to connect to a network. Learning this single PSK could lead to system compromise, and potential alteration of restroom data [17]. Note that the Wi-Fi network used for the IoT buttons was separate from the network used for clinical care. Future iterations of IoT buttons can address this by enabling access to WPA2 Enterprise (also known as 802.11i) that requires a unique username and password and a preinstall unique encryption key, thereby providing additional security, or utilizing onboard cellular networks further isolating the IoT buttons from the hospital network and preventing a system downtime in the event of a Wi-Fi failure [9].

We also experienced human factor challenges in implementation of the IoT button. There are a large number of housekeepers who work in our hospital, and it was difficult to maintain housekeeper adherence to double press the button each time they cleaned a restroom or responded to a request for clean. Our qualitative data reflect that this was an additional step in their workflow; this likely led to a decreased number of double presses confirming cleaning by housekeeping. We therefore were unable to accurately calculate a response time from a request for clean to the time that a housekeeper actually cleaned the restroom. We believe that with continued deployment of the IoT button in our hospital, the process of double presses as the restroom is cleaned will become part of the housekeeping workflow leading to more robust data. Some other potential alternatives to relying upon housekeepers to double press the IoT button include deploying an integrated IoT button and radiofrequency identification (RFID) placard to allow housekeepers to scan their badges when their cleaning is complete, or RFID badges that log the movement of housekeepers into restrooms. This alternative technique will require user acceptability evaluation before the deployment in conjunction with an IoT button system. Despite these challenges, we believe that the IoT button system was able to provide valuable data regarding the pulse of restroom cleanliness in our hospital. During the study period, our housekeeping supervisors perceived the IoT button system as a key part of their daily workflow.

Limitations

This study had several limitations. First, we lacked contextual data surrounding IoT button presses. We do not know if users pressed the button because restrooms were dirty, or in need of service (eg, replenishing supplies and unclogging drains). In

XSL•FO
RenderX

addition, although we were able to record requests for cleaning, we do not know what aspect of the restroom was dirty. Future studies could utilize an in-person monitor to describe reasons for a request for clean and further understand the patterns of the IoT button use. Second, despite acceptability of the IoT button system, we continued to have some human factors issues regarding adoption of a double press to log a completed cleaning procedure of restrooms. We therefore may have underestimated the number of responses to clean restrooms using our system. Third, our qualitative analysis was limited by sample size, and our future research includes an expanded and more rigorous qualitative evaluation on the adoption and sustainability of IoT buttons as part of the housekeeping workflow. Finally, from a technological perspective, the IoT buttons only have a battery life to 2000 presses. A cost analysis of the benefits of replacing IoT buttons in comparison to potential improvements in patient satisfaction may be required to assess the long-term sustainability and value of this technology.

Conclusions

Overall, this study demonstrates that we were able to deploy an IoT button system that measures restroom cleanliness and can also be used by internal staff to log activities related to servicing restrooms. Although there are a variety of technologies that can be used to deliver similar just-in-time notifications regarding restroom cleanliness, we found our system easy to deploy, and engaging. Even 1 year after deployment of the system, we continued to receive notifications through the IoT button for cleaning requests. We additionally were able to understand some potential trends in the ebbs and flows of service request throughout the workday and workweek. We anticipate that with continued optimization of the IoT button system, the data generated could be used to inform alternative staffing models that respond to real-time changes in the need for housekeeping services in a hospital system.

Acknowledgments

The authors thank the staff of the Brigham Digital Innovation Hub and Environmental Services for their support in the execution of this investigation. Visybl Inc initially provided IoT buttons and access to button management interface free of charge. Equipment software administrative tools were later purchased for this study. Visybl Inc was not involved in the drafting or editing of this manuscript.

PRC is funded by NIDA K23DA044874. HS is supported by the following federal grants: AHRQ R01HS024538 (PI: Adelman), AHRQ R01HS024713 (PI: Adelman), NICHD R01HD094793 (PI: Adelman). Work was supported by the Brigham and Women's Hospital Digital Innovation Hub (iHub).

Conflicts of Interest

ABL is a consultant for Abbott on their Medical Device Cybersecurity Council.

References

1. FSR Magazine. Cincinnati: The Harris Poll; 2018. Source: Restroom Reviews Impact Restaurant Visits URL: https://www.foodnewsfeed.com/content/survey-restroom-reviews-impact-restaurant-visits [accessed 2019-06-05] [WebCite Cache ID 78tlB8iP7]
2. Barber N, Scarcelli JM. Clean restrooms: how important are they to restaurant consumers? J Foodserv 2009;20(6):309-320. [doi: 10.1111/j.1748-0159.2009.00155.x]
3. Reeves R, Ramm J, Cornelius V, Graham C. National Health Services. United Kingdom: National Health Services; 2004. Inspecting, Informing, Improving: Patient Survey Report 2004 URL: http://www.nhssurveys.org/Filestore/CQC/IP_KF_2004.pdf [accessed 2019-06-03] [WebCite Cache ID 78rE9GNum]
4. Sofaer S, Crofton C, Goldstein E, Hoy E, Crabb J. What do consumers want to know about the quality of care in hospitals? Health Serv Res 2005 Dec;40(6 Pt 2):2018-2036 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00473.x] [Medline: 16316436]
5. Angelopoulou P, Kangis P, Babis G. Private and public medicine: a comparison of quality perceptions. Int J Health Care Qual Assur Inc Leadersh Health Serv 1998;11(1):14-20. [Medline: 10177366]
6. Barker J, Jones MV. The potential spread of infection caused by aerosol contamination of surfaces after flushing a domestic toilet. J Appl Microbiol 2005;99(2):339-347 [FREE Full text] [doi: 10.1111/j.1365-2672.2005.02610.x] [Medline: 16033465]
7. Knowlton SD, Boles CL, Perencevich EN, Diekema DJ, Nonnenmann MW, CDC Epicenters Program. Bioaerosol concentrations generated from toilet flushing in a hospital-based patient care setting. Antimicrob Resist Infect Control 2018;7:16 [FREE Full text] [doi: 10.1186/s13756-018-0301-9] [Medline: 29423191]
8. Verani M, Bigazzi R, Carducci A. Viral contamination of aerosol and surfaces through toilet use in health care and other settings. Am J Infect Control 2014 Jul;42(7):758-762. [doi: 10.1016/j.ajic.2014.03.026] [Medline: 24818773]
9. Chai PR, Zhang H, Baugh CW, Jambaulikar GD, McCabe JC, Gorman JM, et al. Internet of things buttons for real-time notifications in hospital operations: proposal for hospital implementation. J Med Internet Res 2018 Aug 10;20(8):e251 [FREE Full text] [doi: 10.2196/jmir.9454] [Medline: 30097420]

XSL•FO
RenderX

10. Agency for Healthcare Research and Quality. 2013. Plan-Do-Study-Act (PDSA) Cycle URL: https://innovations.ahrq.gov/qualitytools/plan-do-study-act-pdsa-cycle [accessed 2019-06-03] [WebCite Cache ID 78rFDBkaY]
11. Berwick DM. Developing and testing changes in delivery of care. Ann Intern Med 1998 Apr 15;128(8):651-656. [doi: 10.7326/0003-4819-128-8-199804150-00009]
12. Holden RJ, Karsh BT. 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]
13. R Core Team. The R Foundation for Statistical Computing. 2019. URL: https://www.r-project.org/ [accessed 2019-06-03] [WebCite Cache ID 78rEd1I6a]
14. Michalakis K, Caridakis G. IoT contextual factors on healthcare. Adv Exp Med Biol 2017;989(6):189-200. [doi: 10.1007/978-3-319-57348-9_16] [Medline: 28971427]
15. Blyth PL. Clean restrooms: an issue of quality as well as image. Health Facil Manage 1993;6(8):28, 30, 32-4. [Medline: 10127309]
16. Bandura A. Perceived self-efficacy in cognitive development and functioning. Educ Psychol 1993;28(2):117-148. [doi: 10.1207/s15326985ep2802_3]
17. Chen CY, Hasan M, Mohan S. Securing real-time internet-of-things. Sensors (Basel) 2018 Dec 10;18(12):4356 [FREE Full text] [doi: 10.3390/s18124356] [Medline: 30544673]

Abbreviations

ED: emergency department
IoT: Internet of Things
PSK: preshared key
QR code: quick response code
RFID: radiofrequency identification
SMS: short message service
WPA2: Wi-Fi Protected Access 2

Edited by G Eysenbach; submitted 02.02.19; peer-reviewed by Q Fang, J Gorman, M Whelan; comments to author 28.04.19; revised version received 20.05.19; accepted 20.05.19; published 19.06.19.

Please cite as:
Chai PR, Zhang H, Jambaulikar GD, Boyer EW, Shrestha L, Kitmitto L, Wickner PG, Salmasian H, Landman AB
An Internet of Things Buttons to Measure and Respond to Restroom Cleanliness in a Hospital Setting: Descriptive Study
J Med Internet Res 2019;21(6):e13588
URL: http://www.jmir.org/2019/6/e13588/
doi:10.2196/13588
PMID:31219046

©Peter R Chai, Haipeng Zhang, Guruprasad D Jambaulikar, Edward W Boyer, Labina Shrestha, Loay Kitmitto, Paige G Wickner, Hojjat Salmasian, Adam B Landman. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Tangibility of Personalized 3D-Printed Feedback May Enhance Youths' Physical Activity Awareness, Goal Setting, and Motivation: Intervention Study

Sam Graeme Morgan Crossley[1*], BSc, PhD; Melitta Anne McNarry[1*], BSc, PhD; Parisa Eslambolchilar[2*], BEng, MEng, PhD; Zoe Knowles[3*], PhD, HCPC; Kelly Alexandra Mackintosh[1*], BSc, MSc, PhD

[1]School of Sport and Exercise Sciences, Applied Sports Technology Exercise and Medicine Research Centre, Swansea University, Swansea, United Kingdom

[2]School of Computer Science and Informatics, Human Factors Technology Research Priority Area, Cardiff University, Cardiff, United Kingdom

[3]School of Sport and Exercise Sciences, Physical Activity Exchange, Liverpool John Moores University, Liverpool, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Kelly Alexandra Mackintosh, BSc, MSc, PhD
School of Sport and Exercise Sciences
Applied Sports Technology Exercise and Medicine Research Centre
Swansea University
Academic Office, A110 First Floor
Engineering East, Bay Campus
Swansea, SA1 8EN
United Kingdom
Phone: 44 01792 ext 295075
Email: K.Mackintosh@swansea.ac.uk

Abstract

Background: In the United Kingdom, most youth fail to achieve the government guideline of 60 min of moderate to vigorous physical activity (MVPA) daily. Reasons that are frequently cited for the underachievement of this guideline include (1) a lack of awareness of personal physical activity levels (PALs) and (2) a lack of understanding of what activities and different intensities contribute to daily targets of physical activity (PA). Technological advances have enabled novel ways of representing PA data through personalized tangible three-dimensional (3D) models.

Objective: The purpose of this study was to investigate the efficacy of 3D-printed models to enhance youth awareness and understanding of and motivation to engage in PA.

Methods: A total of 39 primary school children (22 boys; mean age 7.9 [SD 0.3] years) and 58 secondary school adolescents (37 boys; mean age 13.8 [SD 0.3] years) participated in a 7-week fading intervention, whereby participants were given 3D-printed models of their previous week's objectively assessed PALs at 4 time points. Following the receipt of their 3D model, each participant completed a short semistructured video interview (children, 4.5 [SD 1.2] min; adolescents, 2.2 [SD 0.6] min) to assess their PA awareness, understanding, and motivation. Data were transcribed verbatim and thematically analyzed to enable key emergent themes to be further explored and identified.

Results: Analyses revealed that the 3D models enhanced the youths' awareness of and ability to recall and self-evaluate their PA behaviors. By the end of the study, the youths, irrespective of age, were able to correctly identify and relate to the government's PA guideline represented on the models, despite their inability to articulate the government's guideline through time and intensity. Following the fourth 3D model, 72% (71/97) of the youths used the models as a goal-setting strategy, further highlighting such models as a motivational tool to promote PA.

Conclusions: The results suggest that 3D-printed models of PA enhanced the youths' awareness of their PA levels and provided a motivational tool for goal setting, potentially offering a unique strategy for future PA promotion.

(J Med Internet Res 2019;21(6):e12067) doi:10.2196/12067

XSL•FO
RenderX

KEYWORDS

behavior change; health education; feedback; self-monitoring; accelerometry; schools; adolescent; child

Introduction

Background

The government of the United Kingdom recommends that youth (children and adolescents) aged 5 to 18 years should engage in 60 min of moderate-to-vigorous physical activity (MVPA) every day [1] to accrue associated physiological [2,3] and psychosocial health benefits [4,5]. However, only 23% and 20% of boys and girls, respectively, aged 4 to 15 years in the United Kingdom meet these minimum levels of physical activity (PA) [6], with almost 50% of the youths failing to achieve even half the recommended levels [7]. The frequently cited reasons for youth underachievement of the PA guideline are thought to be their lack of awareness of their physical activity levels (PALs) [8-10] and a lack of understanding of what activities and different intensities of PA *count* toward the daily target [11-18]. Given that adults also show a lack of awareness of their PALs [19], have limited knowledge of their respective PA target, and struggle to appropriately identify activity intensities [11], addressing these issues during childhood is important for fostering healthy lifestyle behaviors that can continue into adulthood [20,21].

Based on Weinstein's [22] Precaution Adoption Process Model (PAPM) from the Stages of Change [23], an individual can only be expected to proceed to the contemplation stage when they become aware that their behaviors are not optimal, such as "I do this much MVPA but this much MVPA is recommended" [11]. In a similar way, the Goal Setting Theory [24] notes that setting specific and challenging, yet achievable, goals, in conjunction with feedback regarding performance toward goal attainment, is important to enhance an individual's self-efficacy (ie, an individual's belief to perform a behavior) and health behavior change. In this regard, personalized feedback that represents an individual's PALs in contrast to the recommended level of activity (ie, acting as a goal) is recognized as an important method for raising one's awareness of their PA behaviors and subsequent behavior change [25]. Therefore, for health education to be successful in youths, efforts must be made to first raise an awareness and understanding of their PALs in the form of personalized feedback [8] that supports goal attainment (ie, meeting the recommended guideline) [26]. To make personalized feedback effective, it is important that it is visually stimulating and meaningful to the individual [27,28], as *seeing* makes knowledge credible [29], and greater visibility of feedback contributes to be an added responsibility to act [30,31]. Most personalized feedback is presented through digital on-screen displays (eg, mobile phones or activity tracker displays) [32-38]; however, with recent advancements in three-dimensional (3D) printing technology, Khot et al [39] explored an innovative approach to displaying adults' heart rate data through tangible 3D-printed artifacts to represent a day of PA. This novel approach demonstrated that the visual and tactile nature of the feedback increased adults' awareness of and reflection on their personal PA [39]. Indeed, in youth populations, past research has demonstrated that tangible interfaces can increase youth engagement and reflection in active learning [40,41], with several learning theories placing emphasis on tangibles as tools to stimulate intellectual development in youths [42-44]. Building on these conclusions, more recent formative research has demonstrated that youth have the ability to conceptualize PA data represented as 3D-printed objects [45]. Moreover, 2 age-specific 3D model representations of youth PA data were developed from formative research [45], which were further validated as a potential tool to increase youth awareness and understanding of PA and the recommended guideline [46]. However, the efficacy of the designed age-specific 3D models in a real-world setting as a tool to enhance youth awareness and understanding of PA is currently unknown.

In accord with Forlizzi and Battarbee [47], understanding how a user's experiences change over time in connection to a newly designed product is essential for developing the scalability and potential use of the technology in a realistic context. The user's experience, within the context of technology, is defined by a user's internal state (perceptions, expectations, motivation, and mood), the characteristics of the product (usability, functionality, and purpose), and the context (organizational or social setting) within which the interactions occur with the technology [48]. More recently, video interview methods have become increasingly popular among researchers to assess a user's experiences, understanding, and navigation of newly designed technology [36,49,50]. However, these aforementioned video interviews have either been long in duration (eg, 60 min) [36,49] or been implemented with small numbers of individuals (eg, 16-22 participants) [36,50], which may affect the generalizability of findings.

Objectives

Therefore, the aim of this study was to examine the efficacy of the age-specific 3D-printed models to enhance children and adolescents' levels of awareness and understanding of and motivation for PA during a 7-week faded intervention, whereby the youths receive personalized 3D-printed models displaying their PALs. It is hypothesized that receiving personalized 3D-printed PA feedback will enhance the youths' (1) awareness of their MVPA levels compared with the government guideline of 60 min of MVPA; (2) understanding of what constitutes PA and of moderate-and-vigorous-intensity activity; and (3) motivation to be more physically active.

Methods

Participants

A total of 2 primary schools and 1 secondary school in South Wales, United Kingdom, were invited to participate in the intervention study. In total, 97 youths participated in the study, of which 39 were primary school children (22 boys; mean age 7.9 [SD 0.3] years) and 58 secondary school adolescents (37 boys; mean age 13.8 [SD 0.3] years). All primary school children were white British, with 96% of secondary school

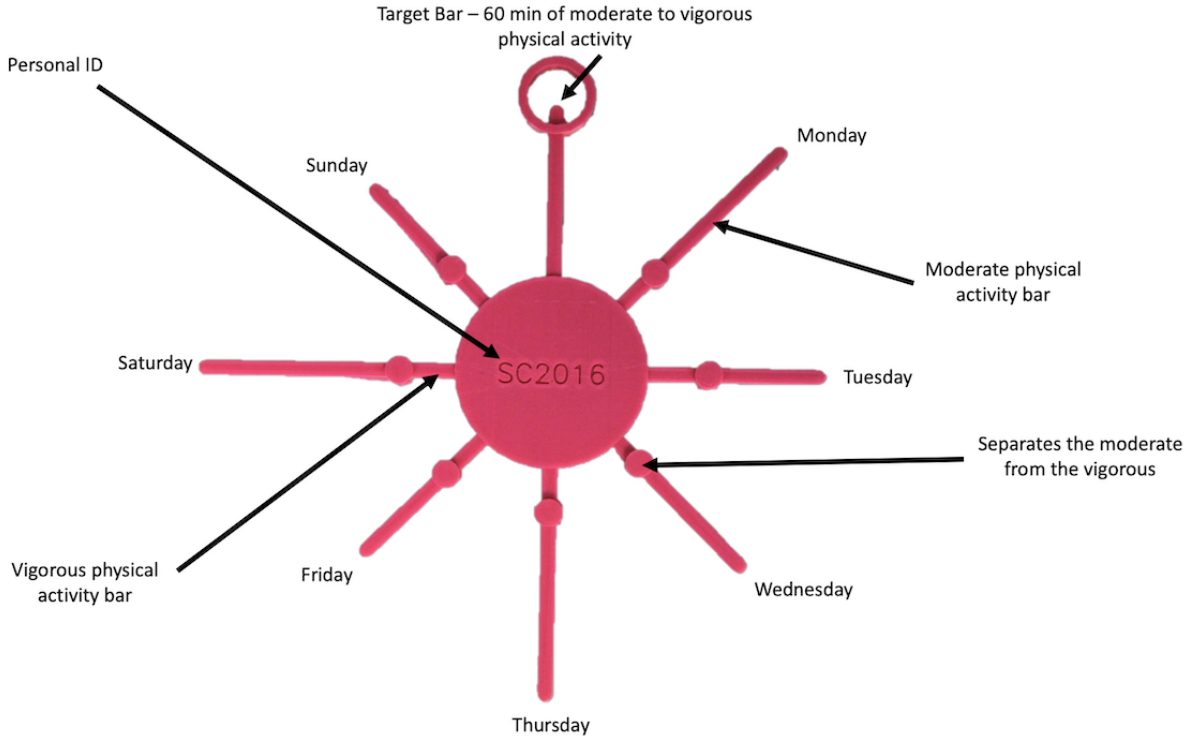
XSL•FO
RenderX

adolescents being white British and the remaining 4% being Asian (2%; n=1) and black British (2%; n=1). All participants returned informed parental or carer consent and child assent before participation. Ethical approval was granted by the University Ethics Committee and conducted in accordance with the Declaration of Helsinki (ref: PG/2014/40).

Intervention Design

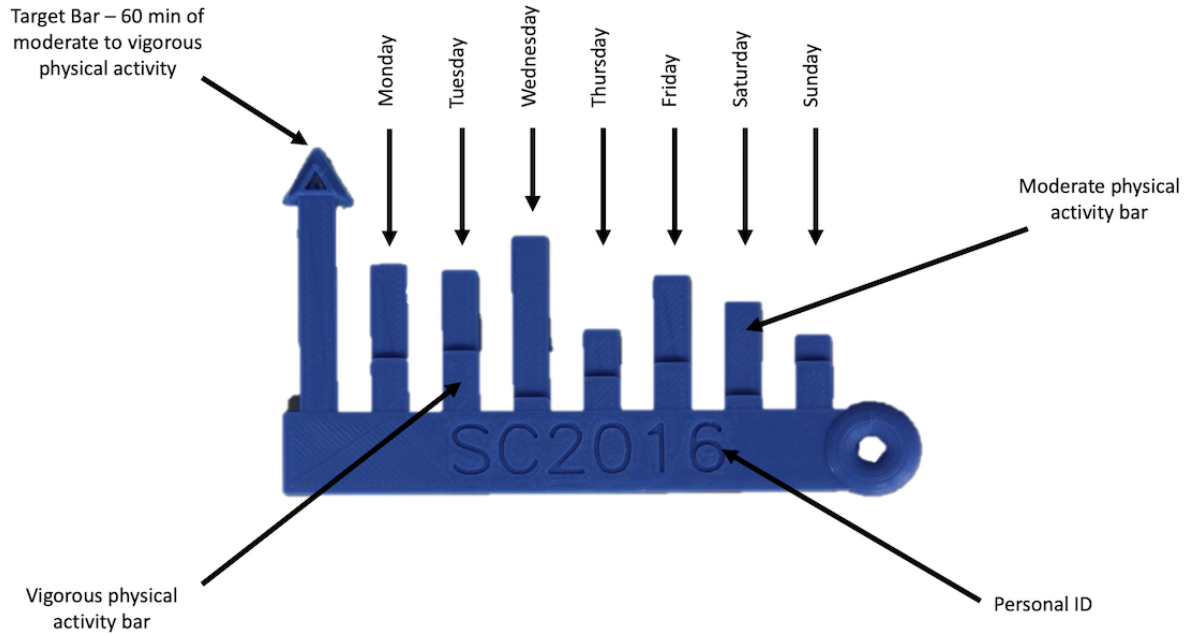
The 3D-printing PA intervention was informed by 2 previous user-centered, qualitative approaches that explored the needs, preferences for content and designs, and understanding of 3D-printed models among the youths (ie, children and adolescents) as described in detail elsewhere [45,46]. To encourage lifestyle change, the intervention was theoretically based, in part, on the notion of youths being visual and tactile learners [42,51,52], with an emphasis on the PAPM [22] and Goal Setting Theory [24] as ideologies to enhance awareness of behaviors in relation to set goals through personalized feedback that encompasses a physical incentive. The intervention was implemented for 7 weeks to align with the school term time. The intervention was designed to objectively measure the youths' weekly PALs and use these data to generate personalized age-specific 3D-printed models to represent the amount of moderate PA and vigorous PA achieved each day across a week as well as display the PA guideline of 60 min of MVPA (Figures 1 and 2). The intervention employed a novel approach that involved participants receiving a total of 4 age-specific 3D-printed models over the course of the 7-week intervention in a faded manner. Specifically, the youths received their 3D models following baseline (model 1=M1), week 1 (model 2=M2), week 3 (model 3=M3), and after week 6 (model 4=M4). The faded approach has been proposed as a method for maximizing the long-term effectiveness of feedback compared with frequent feedback, which only provides short-term benefits [53]. In this regard, the faded method is underpinned by starting with high levels of feedback and then, as the participant begins to master the components of the task, gradually reduce or fade the feedback until the person is performing the task autonomously [54-58]. A key point to this faded design is to increase the sustainability and real-world *implementability* of 3D-printing PA interventions by examining how the 3D models can be integrated into the everyday lives of youths to determine the success of deployment and adoption of the models [59]. Participants received their personal 3D-printed model 1 to 3 days after PA measurement. Immediately following the receipt of each 3D model, all participants completed an individual, semistructured short video interview conducted by the first author either during their physical education class (secondary school) or in an appropriate quiet area within the school environment (primary school) to elicit information on study outcomes [60]. Video interviews are considered a viable method for recording the youths' experiences with technological designs [50]. All participants received 1 instruction manual (Figures 1 and 2) for their respective age-specific 3D model after completing their first short individual interview to obtain baseline perceptions of primary outcome measures.

Figure 1. Children's Sun age-specific three-dimensional model of physical activity instruction manual. 3D: three-dimensional; PA: physical activity.



XSL•FO
RenderX

Figure 2. Adolescents' bar chart age-specific of physical activity instruction manual. PA: physical activity.



Procedures

Anthropometrics

All the participants' standing stature, body mass, and waist circumference were measured according to the techniques outlined by the International Society for the Advancement of Kinanthropometry [61]. Participants were required to be in minimal clothing (ie, shorts and T-shirt) and barefoot. Body mass was measured to the nearest 0.1 kg using an electronic weighing scale (Seca 876), with stature assessed to the nearest 0.1 cm using a portable stadiometer (Holtain Sitting Height Stadiometer, Holtain Ltd). Body mass index (BMI) and weight status were calculated from stature and body mass measurements as a proxy for adiposity [62]. Based on BMI z-score calculations, age- and sex-specific BMI cut-points from the United Kingdom were applied to categorize participants as underweight, normal weight, or overweight/obese [63]. All anthropometric measurements were conducted within the school by trained research assistants under the supervision of the first author.

Measuring and 3D Printing of Physical Activity Data

All participants were asked to wear the wGT3X-BT triaxial accelerometer (ActiGraph LLC) on an elastic belt positioned on their right midaxilla line at the level of the iliac crest for 7 consecutive days to provide an objective estimate of their PALs. Numerous studies have validated the wGT3X-BT triaxial accelerometer as a valid and reliable objective measurement of the quantity and frequency of PA [64-66], with previous research demonstrating that the hip placement is the most precise single location to detect everyday activities [66,67]. All participants were shown a demonstration of the accelerometer hip placement via Sam Graeme Morgan Crossley and provided an information sheet regarding the use and safety of the device at baseline measurement. As far as practically possible, participants wore the same accelerometer (serial number) at each time point to remove *between-unit* variation [68]. Participants were instructed to wear the accelerometer all the time (24 hours per day), except for when engaging in water-based activities (swimming, showering, and bathing) and contact sports. Accelerometers were activated to run from midnight the day participants received the accelerometer until midnight 7 days later and initialized to record raw accelerations at a frequency of 100 Hz. Following the collection of accelerometers, participants' 7-day PA data were then downloaded and analyzed using Actilife version 6.13.3 (ActiGraph LLC). Given the intervention was designed to provide all participants with raw feedback on MVPA levels, even if the accelerometer was removed (eg, for water-based or contact sport activities), no inclusion criteria were applied to the accelerometry data. Therefore, implications for the youths not wearing the accelerometer on one or more days would result in them receiving a 3D model with no data displayed on that specific day. Each day's MVPA level was calculated using Evenson child cut-points [69], which are known for providing the closest estimates of moderate-and-vigorous-intensity PALs during the free-living measurement [70]. Participants' MVPA levels and personal ID code (eg, participant initials and model number) to distinguish participants' personal age-specific 3D model were then inserted into the age-specific custom-developed 3D model code loaded on OpenJSCAD version 1.8.0 and subsequently 3D-printed using ABSplus filament on the Objet 1000 (Statasys).

To examine participants' baseline MVPA, data were further analyzed using KineSoft (version 3.3.67; KineSoft), employing 1-second epochs with sustained periods of at least 20 min at zero counts considered as nonwear time [71]. Participants were included in the analysis if they met the minimum daily wear-time criteria of 10 hours for any 3 days [72], which has previously been shown to produce reliable estimates of PA in

XSL•FO
RenderX

youth [73]. PA intensities were calculated using cut-points in a study by Evenson et al [69]. Data collection took place during the school term from January to April 2017; therefore, PA data were representative of usual winter or spring free-living activities.

Short Individual Video Interviews

Short, individual interviews were chosen as they lend greater control to the interviewer over the interview process relative to the unpredictable nature of focus group interactions [74]. Individual interviews also allow the researcher to locate specific ideologies within particular individuals [75], which is not always possible within focus groups given that the youths may tag onto the views of others without necessarily reflecting on the value or meaning [76]. To reinforce the interpretation of the qualitative data, each individual interview was filmed to capture the youths' nonverbal and contextual understandings of the 3D model that could be missed in a narrative statement alone [77]. The interviews were semistructured so that the facilitator could ask probing questions around the predefined topics and to keep discussions relevant to the study aims [78]. The 2 interview types (children and adolescents) were conducted using the same research protocol and followed a predefined schedule of questions (Table 1) that sought to address concepts on youth, awareness of the youths' PALs, understanding of intensities and interpretations of the 3D model, and motivational benefits and utility of the 3D models. A total of 369 interviews were digitally voice-recorded (Olympus DM-520 digital voice recorder) and video-recorded (Sony Handycam HDR-PJ540), lasting 4.5 [SD 1.2] and 2.2 [SD 0.6] min, for children and adolescents, respectively. All interviews were transcribed verbatim, resulting in 816 pages (386 and 430 pages for children and adolescents, respectively) of raw transcription data, Arial font size 12, and double spaced.

Table 1. Example interview questions.

Topic	Examples
Motivation or awareness of PALs[a]	What do you think of your first 3D[b] model?
Understanding of PA[c]	What you think physical activity means?
Awareness of PAL or understanding of model	How does your 3D-printed model show your physical activity?
Understanding of intensity	What kind of activities might be vigorous and moderate physical activities?
Motivation or model utility	What will you do with your 3D model now?

[a]PALs: physical activity levels.

[b]3D: three-dimensional.

[c]PA: physical activity.

Data Analysis

A Shapiro-Wilks test was used to confirm data normality within the anthropometric and PA data. Once normal distributions were confirmed, independent sample *t* tests were used to assess differences between sexes within children and adolescents. All statistical analyses were conducted using IBM SPSS Statistics 22 (SPSS Inc), and statistical differences were accepted at $P \leq .05$. Through the process of content analysis, transcripts were approached qualitatively to focus on the context of the youths' awareness of their PALs and preunderstanding of intensities and the motivational aspects of the 3D models. To quantify patterns within the different time points (ie, receiving model 1-4), it was quantitatively noted as to the number of participants who were associated with specific statements and for the classification of categorical data being accurate (ie, correct interpretations of the 3D model and activity intensities) [79]. To aid and align the accurate classification of 3D model interpretations, interview videos were also assessed to examine participants' nonverbal interactions with their 3D model by noting gestures (eg, correctly points to the 60-min MVPA guideline bar) within transcripts [77]. All transcripts were thematically analyzed by the first author, first by data immersion, which involved *repeated reading* of the transcripts in an active way, searching and noting of meanings and patterns within the dataset [80]. Following the initial data immersion process, coding was undertaken, using a manual cut and paste technique, which allowed for the data to be organized into groups that were considered pertinent to the research questions. All codes were then sorted into potential themes by collating all relevant coded data extracts to the newly identified theme. The frequency counts and themes with indicative quotes were then represented diagrammatically using a pen profile approach [14,81-83], with the percentages of youth expressing specific themes calculated from frequency counts. The first author discussed the identified themes with the last author to determine the existence of relationships within the data. Themes that did not have enough supportive data or were too diverse were discarded. The second author critically cross-examined the data through reverse triangulation, from the pen profiles back to the transcripts, until all alternative interpretations of the data were exhausted. The pen profiles were then critically reviewed by all other authors, allowing further interpretations of the data until a final consensus was reached.

Results

Descriptive Information

Participants' anthropometric characteristics and PALs are displayed in Table 2. There were no significant sex differences between children, but adolescent boys were significantly taller and heavier than their female counterparts. At baseline, 13% (5/39; boys, 13%, 3/22; girls, 12%, 2/17) of children were overweight or obese with the remaining 87% (34/39; boys, 87%, 19/22; girls, 88%, 15/17) of children being classified as normal

XSL•FO
RenderX

weight, with no children being underweight. For adolescents, 22% (13/58; boys, 16%, 6/37; girls, 33%, 7/21) were overweight or obese, and 78% (49/58; boys, 84%, 31/37; girls, 67%, 14/21) were normal weight, with no individuals categorized as underweight. Valid baseline accelerometer data were collected from 68% (66/97) of the consenting participants, with 72% of both children (28/39) and adolescents (42/58) meeting the wear-time criteria. Irrespective of age, there were no significant differences between sexes for baseline MVPA. The provision of baseline MVPA data showed that only 38% (15/39) and 26% (15/58) of children and adolescents, respectively, met the recommended daily MVPA guideline.

Table 2. Descriptive and anthropometric characteristics of participants.

Characteristics	Primary			Secondary		
	Boys (n=22)	Girls (n=17)	Both (n=39)	Boys (n=37)	Girls (n=21)	Both (n=58)
Age (years), mean (SD)	7.9 (0.3)	7.8 (0.35)	7.9 (0.3)	13.8 (0.3)	13.7 (0.3)	13.8 (0.3)
Stature (m), mean (SD)	1.28 (0.1)	1.25 (0.1)	1.27 (0.1)	1.66 (0.1[a])	1.63 (0.1)	1.65 (0.1)
Waist circumference (cm), mean (SD)	58.1 (4.9)	59.6 (5.1)	58.7 (5.0)	73.3 (6.0)	69.2 (6.3)	72.1 (6.4)
Body mass (kg), mean (SD)	26.1 (3.5)	25.8 (4.0)	26.01 (3.5)	56.05 (10.2[a])	55.8 (6.8)	55.9 (9.0)
BMI[b] (kg/m^2), mean (SD)	15.9 (2.0)	16.6 (2.4)	16.2 (2.03)	20.2 (2.4)	21.1 (3.0)	20.55 (2.7)
Weight status n (%)						
Underweight	__[c]	—	—	—	—	—
Normal weight	19 (87)	15 (88)	34 (87)	31 (84)	14 (67)	49 (78)
Overweight/obese	3 (13)	2 (12)	5 (13)	6 (16)	7 (33)	13 (22)
Physical activity levels						
Baseline MVPA[d] (min[e]), mean (SD)	63.3 (11.8)	63.3 (13.4)	63.3 (12.3)	57. 4 (15.4)	50.1 (15.0)	54.6 (15.5)
MVPA guidelines, n (%)	8 (36)	7 (32)	15 (38)	12 (32)	3 (14)	15 (26)

[a]Significant difference between boys and girls within an age group (P<.05).

[b]BMI: body mass index.

[c]Not applicable.

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

[e]Numbers represent provision of valid days (3 valid days with a minimum 10 hours wear time).

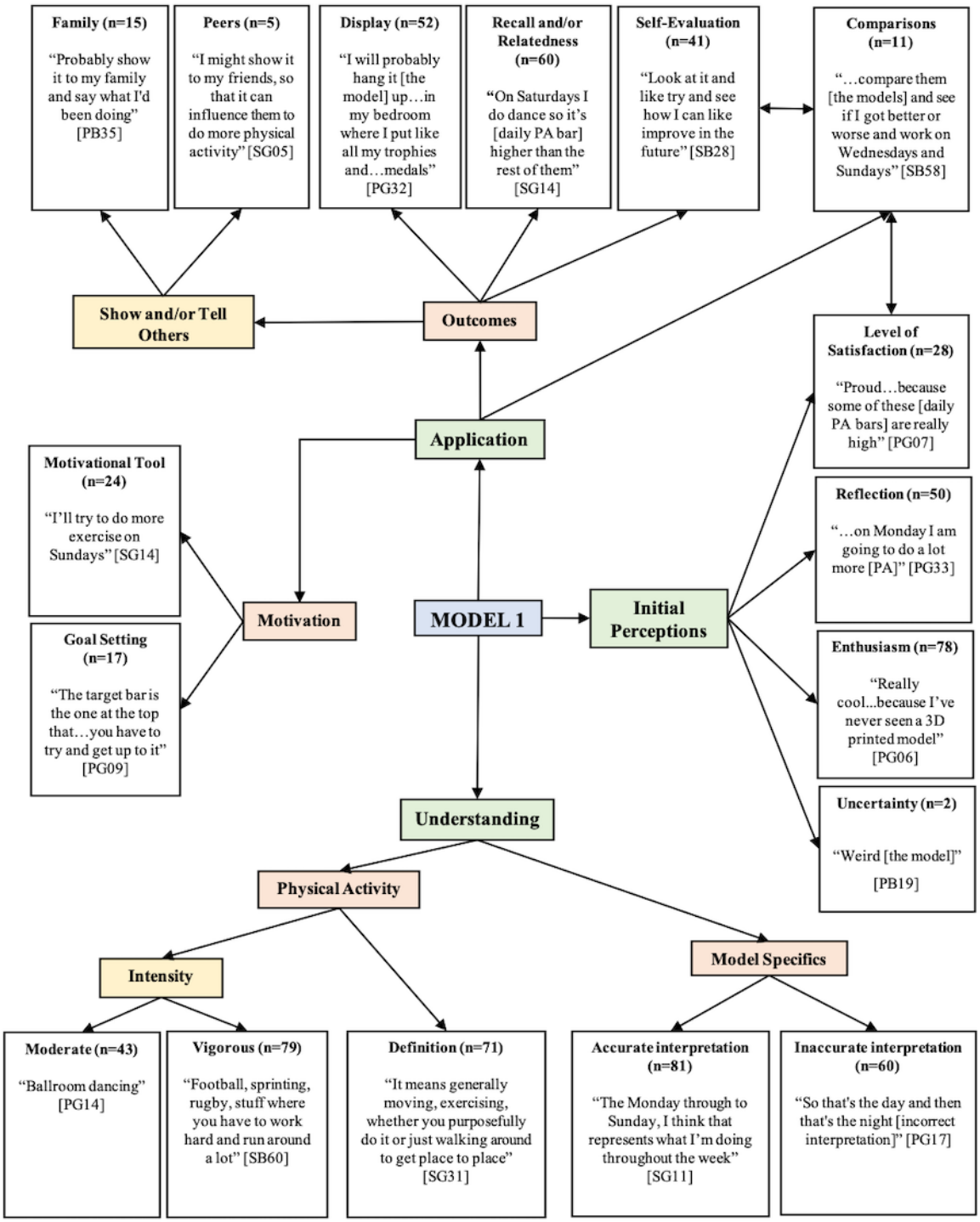
Primary Outcomes

The first model outcomes for children's and adolescent's data are combined and presented in 1 pen profile (Figure 3), as no different themes were found from independent analyses. To avoid duplicating the pen profiles and their identified key themes, Table 3 displays the youths' frequency of occurrence of key themes for each of the 4 3D models, with children, adolescents, and sex independently split.

Following the first model, the majority of youths (78/97, 80%) expressed a high level of enthusiasm for their 3D model, expressing that it is "really cool...because I've never seen a 3D-printed model" (PG06, M1). However, by the final model, only 4% (4/39) of children and no adolescents still expressed similar enthusiasm. Despite this, 28% (28/97) of youths displayed satisfaction on how they were "very proud [of the model]" (PG07, M1) of their first 3D model, with this level of satisfaction increasing from 39% (38/97) to 60% (58/97) and 68% (66/97), by the second, third, and fourth models, respectively. Furthermore, the youths demonstrated increased levels of reflection through the 3D models upon how they "...never thought Saturday was going to be that long" (PB35, M3), from 51% (50/97) to 60% (62/97) and 66% (64/07), for the first, second, and third models, respectively, although by the fourth model, this level of reflection dropped to 58% (56/97).

XSL•FO
RenderX

Figure 3. Youths' pen profile Model 1. 3D: three-dimensional; PA: physical activity.



XSL•FO
RenderX

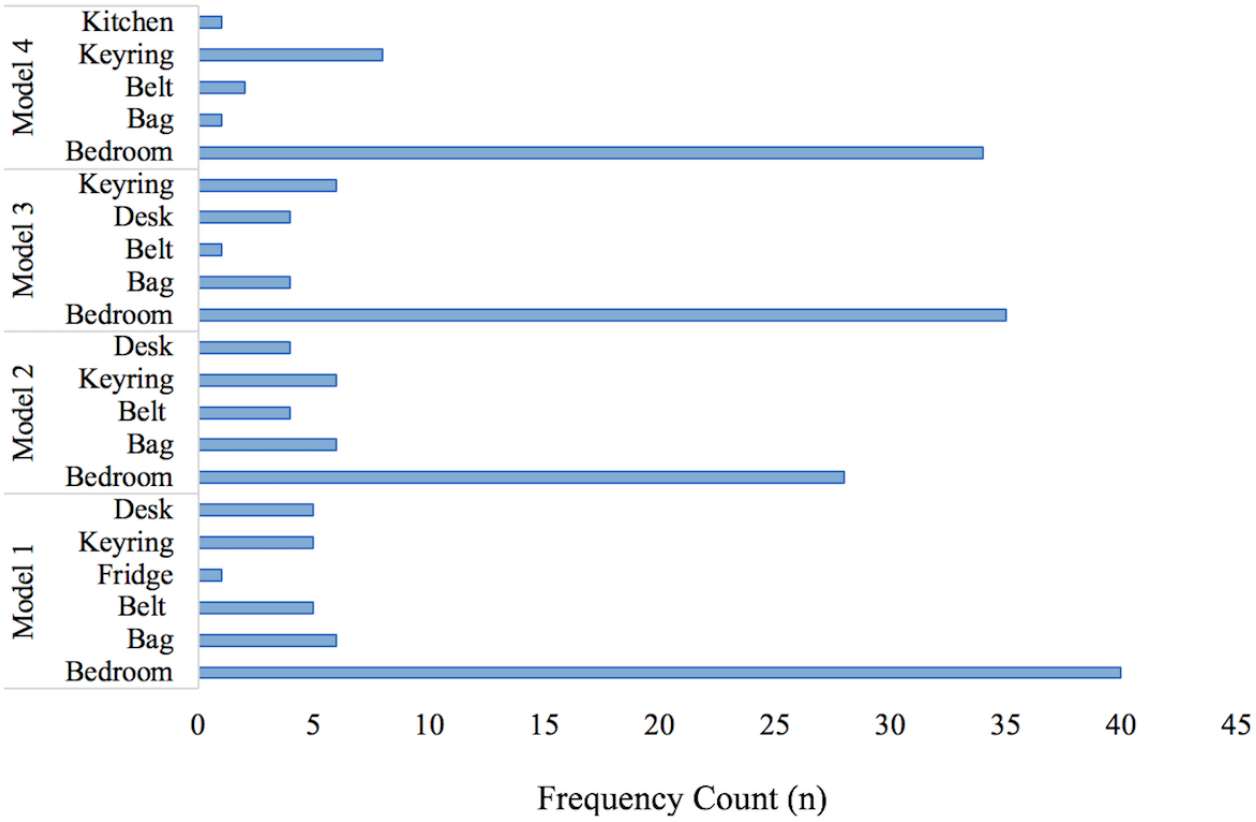
Table 3. The youths' frequency of occurrence of key themes (n=97). n indicates frequency counts, and superscripted G indicates Girl frequency count.

Themes	Model 1 (n^G)			Model 2 (n^G)			Model 3 (n^G)			Model 4 (n^G)		
	Children	Adolescents	Total	Children	Adolescents	Total	Children	Adolescents	Total	Children	Adolescents	Total
Enthusiasm	30^{15}	48^{19}	78	15^{4}	4^{2}	19	9^{3}	0	9	4^{1}	0	4
Level of satisfaction	8^{3}	20^{6}	28	10^{2}	28^{14}	38	20^{10}	38^{16}	58	24^{9}	42^{17}	66
Reflection	13^{6}	37^{12}	50	24^{7}	38^{18}	62	20^{9}	44^{15}	64	14^{6}	42^{16}	56
Uncertainty	2	0	2	2	0	2	0	0	0	0	0	0
Definition	26^{7}	45^{18}	71	27^{9}	41^{14}	68	32^{11}	42^{17}	74	30^{11}	44^{15}	74
Moderate intensity	5^{2}	38^{15}	43	15^{5}	34^{12}	49	9^{6}	36^{15}	45	12^{2}	34^{14}	46
Vigorous intensity	29^{10}	50^{19}	79	31^{10}	47^{17}	78	33^{12}	49^{20}	82	32^{13}	47^{20}	79
Accurate interpretation	34^{14}	47^{18}	81	36^{14}	46^{17}	82	35^{14}	51^{21}	86	36^{14}	49^{19}	85
Inaccurate interpretation	33^{13}	27^{9}	60	19^{6}	27^{9}	46	21^{9}	29^{10}	50	18^{7}	32^{11}	50
Comparisons	1	10^{3}	11	20^{8}	27^{10}	47	20^{6}	30^{13}	50	14^{4}	28^{14}	42
Goal setting	10^{2}	7^{3}	17	21^{9}	24^{8}	45	31^{13}	43^{16}	74	31^{14}	40^{15}	71
Motivational tool	5	19^{7}	24	11^{2}	19^{10}	30	18^{7}	22^{11}	40	11^{2}	23^{9}	34
Recall and/or relatedness	18^{6}	42^{15}	60	19^{7}	38^{7}	57	28^{12}	40^{16}	68	30^{12}	43^{17}	73
Self-evaluation	11^{4}	30^{11}	41	20^{6}	39^{15}	59	22^{9}	44^{18}	66	18^{6}	39^{17}	57
Display	23^{10}	29^{10}	52	25^{10}	21^{5}	46	28^{14}	23^{7}	51	24^{12}	20^{8}	44
Family	8^{4}	7^{4}	15	5^{2}	4^{4}	9	5^{2}	7^{6}	12	3	5^{4}	8
Peers	3	2^{2}	5	1	3^{1}	4	1^{1}	2^{1}	3	1	0	1

Overall, the youths showed little difference in their interpretations of their meaning for PA (M1, 71/97, 73% to M4, 74/97, 76%), stating it is "like doing sports and stuff that includes moving your body" (PB20, M2), with similar outcomes on their interpretations of the intensities of moderate (M1, 43/97, 44% to M4, 46/97, 47%) "like walking" (SB55, M2) and vigorous (M1, 79/97, 81% to M4, 79/97, 81%) "like sprinting so your heart rate is like beating at a fast pace" (SB45, M3). Moreover, across all time points, only 5% (2/39) of children and 17% (10/58) of adolescents were able to relate the guideline bar accurately to "60 minutes of exercise a day" (SG42, M3), with only a small proportion of adolescents (3/58, 5%) able to articulate the guideline of "...at least an hour of hard and moderate activity every day" (SB49, M3). However, the youths demonstrated an accurate ability to interpret the basic components of the 3D models (eg, days and high and low PALs) from the first (81/97, 83%) to the fourth model (85/97, 88%), such as "It [the model] means the days of the week and how much activity you've been doing" (PG31, M1), with adolescents being able to correctly distinguish "this one [vigorous bar] is the high-intensity sport activities and this one [moderate bar] is the more moderate sport activities" (SB03, M4). Moreover, the youths were able to correctly interpret and identify with "the target bar...that shows how much exercise you should do in a day, which is one hour" (PB10, M3). As a consequence, the youths increasingly adopted the guideline bar as a goal-setting strategy, from the first (17/97, 18%) to the second (45/97, 46%) and third models (74/97, 76%), with a small drop following the fourth model (71/97, 73%). Specifically, the youths demonstrated this goal setting through how "I haven't reached the target point [on] Monday" (SG09, M3) and "you have to try and be higher than that arrow [guideline bar] and that would be you reaching your target" (SG35, M4). Conversely, some youths expressed inaccurate interpretations of their 3D models; however, this number dropped with time from the first (60/97, 62%) to the fourth model (50/97, 52%). Of note were the small number of children (10/38, 26%) by the final model who were able to correctly interpret the moderate and vigorous bar representations, with children most commonly mistaking the bar as "the morning [vigorous bar] and that's the afternoon [moderate bar]" (PB08, M3). For adolescents, only 14% (8/58) demonstrated to incorrectly identify "the lower solid bar [vigorous bar] is walking activity, and the higher bar [moderate bar] is like sprinting activity" (SB52, M4).

XSL•FO
RenderX

Figure 4. Youths display preferences for three-dimensional models. 3D: three-dimensional.



In terms of the application of the 3D models, 11% (11/97) of the youths expressed that they would "compare the next one [3D model] with it [the current model], and I'll try to do more exercise on Sundays" (SG14, M1), with this application of the models increasing following the second model (47/97, 48%), with no substantial change for time points thereafter. From the first model, 42% (41/97) of the youths demonstrated self-evaluation of their PALs on how "I need to improve certain days and do more on certain days than others" (SG32, M1), with this self-evaluation increasing from 61% (59/97) to 68% (66/97) for the second and third models, respectively. Interestingly, a higher percentage (18/21, 86%) of adolescent girls appeared to self-evaluate their PALs, expressing they would "see if there's anything I can change to get a higher activity than what I got" (SG37, M2).

Throughout all time points (M1, 60/97, 62%; M2, 57/97, 59%; M3, 68/97, 70%; M4, 73/97, 75%), there was little change in the youths' ability to recall and relate their 3D models to their past week of PA, expressing how "on Saturdays I do dance so it's bar of activity is higher than the rest of them" (SG14, M1). Some youth reported the use of the 3D models as a motivational tool because "it's [the 3D model] kind of encouraging me to do more activity, so I can get the bar higher [on the 3D model]" (SB19, M2), with this perception increasing from 25% (24/97) to 31% (30/97) to 41% (40/97) for the first, second, and third models, respectively. From all time points, only 5% (5/97) of the youths expressed that they would "show it [the 3D model] to my friends" (PB01, M1), with a larger number of the youths (11/97, 11%), of which were highly representative of adolescent girls and children of both sexes, expressing how they would "probably like show my parents the model" (SG43, M3). Almost half the number of the youths (47/97, 48%) mentioned that they would display their 3D models in their house (Figure 4), with this proportion slightly greater in children, with a preference to "hang the model up in my bedroom" (PB11, M2).

Discussion

Principal Findings

The primary aim of this study was to evaluate the effectiveness of 3D models of PA to enhance youth awareness and understanding of and motivation to engage in PA. Taken together, the findings suggest that the 3D model feedback offered a unique strategy to enhance youth awareness of their PALs and associations to the government guideline as well as provide the youths with a motivational tool for goal setting.

In this study, 63% (62/97) of the youths demonstrated that they were able to quickly interpret the basic components of their first 3D model (eg, the different days of activity and their low and high PALs). Indeed, these initial interpretations of the age-specific 3D models are promising given that previous research highlights that being able to quickly interact and interpret a tool, such as 3D model, enables an individual to learn about their behaviors from the start, all of which makes the experience with the tool rewarding and minimizes the potential for abandonment [47]. Following the receipt of their final 3D model, 59% (57/97) of the youths self-evaluated how the 3D models had made them "more aware" (SB58, M2) of their PALs. It could be argued that this raised awareness was a direct result of wearing the accelerometer rather than the 3D model per se;

XSL•FO
RenderX

however, this is unlikely, as evidence suggests that accelerometers alone do not develop youth awareness of PA [84]. A more likely reason for this increased awareness may have been the utilization of an objective measure of PA in combination with personalized feedback, which has previously been suggested as an effective combination to raise an individual's awareness of their PA [85]. Complementary to this understanding, the PAPM [22], from the Stages of Change [23], suggests that an individual is unlikely to proceed to the contemplation stage unless they become aware that their behaviors are inadequate. Base on this notion, this study demonstrated that 68% (27/39) and 78% (45/97) of children and adolescents, respectively, were able to identify that "some days I'm reaching the guideline bar, but some days I need to do more physical activity" (SB51, M4). This ability to apply their respective 3D model guideline bar to their personal PALs is a positive indicator, given that previous research has shown that the youths who are aware of their PALs and the recommended guideline are on average 20 min more active than their unaware counterparts and, consequently, more likely to achieve the 60 min of MVPA [8,86-88]. Therefore, given that awareness of risk behaviors is identified as an independent correlate for behavior change [89], the 3D-printed feedback may not only be important to help the youths categorize themselves into the correct stage of change (ie, precontemplation, contemplation, and preparation=not meeting the guideline, vs action and maintenance=meeting the guideline) but also to help the youths perceive the need to change behavior [90], warranting further investigation.

One important consideration with regard to 3D-printed feedback is that it possesses a higher level of visibility compared with digital feedback (ie, on a mobile phone) within the physical world [91]. In this way, 3D-printed PA data are more publicly visible to peers, teachers, and family members. In contrast to previous perceptions [45], only 5% (5/97) and 11% (11/97) of the youths in this study reported that they compared their models with their peers' models and showed their family members the models, respectively. Despite this, it could be speculated from previous research that the youths may have more frequently compared their 3D models with those of peers within the playground and classroom environments [92]. Moreover, it is also likely that family members did indeed come into regular contact with the 3D models, given the range of ways that the youths (53/97, 54%) displayed their models in the bedroom, on their school bag, or attached to the house keys. In this regard, it is important to consider how the visibility of the 3D models may have stimulated more social interactions with peers and family and thus influenced the youths' levels of self-evaluation (58/97, 59%, M4) and reflection (57/97, 58%, M4) of their PALs, rather than the 3D model itself. Indeed, the involvement of peers [93-95] and family [94,96-99] can play a significant role in motivating the youths to be more engaged in PA. On the contrary, sharing and comparing 3D models with peers may induce a competitive environment, which can lead to negative feelings of the self and peer pressure to engage in an activity [100]. Of concern are adolescent girls, who have been shown to be particularly vulnerable at this age to body dissatisfaction, a time when self-awareness, self-consciousness, and preoccupation with self-image all dramatically increase [101]. Indeed, a number of adolescent girls (9/21, 43%) in this study reflected on how being perceived as physically active according to the 3D model was important because "you'll be more confident because like people won't judge you" (SG34, M1) and worried about how "if you're not active you'll end up having a very, well kind of not nice figure [*body shape*]" (SG14, M2). As a consequence, the youths who display such feelings of pressure and guilt for not achieving enough PA may remove themselves from engaging in peer comparisons (eg, sharing their PALs with others) altogether [100,102] and abandon the 3D model. These issues do question how public displays of PA data could intrude upon an individual's privacy [103]. In this light, future research should look to monitor more closely how youth, and in particular adolescent girls, personally reflect and evaluate their PALs with regard to body image and the influence of interactions and support from significant others on PALs.

Following receipt of the final 3D model, 72% (71/97) of the youths had seemingly adopted the guideline bar as a goal-setting strategy, expressing how they monitored their goal-related progress through the guideline bar represented on the 3D models. In this way, the 3D models acted as an important moderator for participant goal attainment, which subsequently led to the youths' self-determined adjustment of PA strategies (eg, starting to play football) and/or effort levels (eg, try harder to do more exercise) [104,105]. As noted within the Goal Setting Theory [24], and addressed in the Social Cognitive Theory [106], setting specific and challenging (yet achievable) goals with feedback on goal attainment is an important step to enhancing an individual's self-efficacy (ie, their belief to carry out a behavior) and thus, behavior change. Numerous reviews support the effectiveness of goal setting to promote the youths' PA engagement [107-109], whereas others suggest that feedback alone has a motivating effect, regardless of whether the feedback is tied to a specific goal or not [110-112].

One particular dimension of the Goal Setting Theory [24] that resonates with this study's findings is the notion that goal attainment can be enhanced by incorporating feedback with rewards (eg, monetary rewards that are linked to goal achievement). Indeed, throughout the intervention, 57% (56/97) of the youths expressed how they would display their 3D model in their bedroom, with some revealing how they placed their models next to their prized "trophies and medals" (PG32, M1). In this way, it could be argued that 3D-printed feedback is received by the youths as a reward of their PA achievements, which is known to heighten an individual's success toward a goal as opposed to just setting a goal alone [113]. According to Locke and Latham [114], rewards are important to sustain a person's interest in PA, which may stand true given the success of incentive-based interventions in promoting PA of youth [115-117]. On the contrary, it is important to consider the influence of a reward or incentive on youths' intrinsic interest to engage in PA as an explicit means to receiving the extrinsic reward (eg, 3D model) and, once removed, whether their behavior reverts back to baseline [118,119]. However, a recent systematic review provides strong evidence that behavioral incentives are an effective means of encouraging PA in youth, suggesting that there is a wide range of incentive designs that are yet to be explored [120]. Perhaps the novelty of 3D printing

XSL•FO
RenderX

PA feedback may offer a greater learning value than previous incentive-based designs, as a result of the 3D models being a blend of a reward, feedback, and goal attainment that embodies personalized data and represents the active self [103]. Therefore, this study supports the utilization of tangible feedback as a novel goal-setting strategy for PA of youth through a reward, feedback, and goal attainment, each of which are known to elicit greater self-efficacy [106,113] and youth engagement within interventions [121]. However, for the aforementioned conclusions to be credible, future research should seek to account for youth MVPA levels in response to the tangible representations of PA guidelines (ie, 60 min of MVPA daily) or personal goals as a tool to support youth engagement and understanding of their PA behaviors.

On the basis of previous *learning styles* that support the use of tangibles to inform intellectual development and enable higher mental functions in youth [40,42-44,122], it was originally postulated that the present 3D-printed feedback of PA may enhance the youths' comprehension of intensities (ie, MVPA) and associations with the government guideline [46]. However, only 5% (2/39) of children and 17% (10/58) of adolescents, across all time points, were able to interpret the guideline bar in terms of the number of minutes (ie, 60 min), whereas no children and 5% (3/58) of adolescents were able to cite "1 hour of physical activity whether it's moderate or vigorous" (SB60, M4). These findings align with previous research suggesting that particularly children have a lack of ability to define time [123,124] and intensity in the context of PA [12,13,15-17]. Indeed, these findings fuel the present debate to whether *learning styles*, such as youth being *visual and tactile* learners [42], are effective strategies to enhance an individual's understanding of information [125]. Previous research has demonstrated that changing the learning mode or strategy for a specific population had little improvement on learning outcomes to justify the time and financial costs involved [126-129]. Therefore, the study's findings question the use of tangibles as an effective means to enhance the youths' comprehension of the MVPA terms associated with the guidelines. Future research may wish to explore different 3D model designs using inscriptions of the intensities, moderate and vigorous, on the 3D models to aid the youths' comprehension of terms.

There are a number of inherent challenges associated with 3D printers and their slow development process that should be noted. Specifically, the process of creating the 3D models, following the downloading, analyzing, and mapping of the youths' PA data onto the 3D models and subsequent 3D printing, involved a considerable amount of time, which consequently delayed the delivery of feedback to the youths. It could be speculated that this delayed timing of the feedback may have negatively affected youth awareness of their PA behaviors, especially given the evidence of the youths' limited capacity to recall their previous activities [130-132]. Adding to this issue was the number of youths who played contact activities, which involved "taking it [the accelerometer] off because of rugby training" (SB24, M3) and, consequently, "forgetting to put it [the accelerometer] back on again" (SB25, M3). In this regard, the 3D models did not account for PA in the form of water-based activities and contact sports, which are likely to contribute to daily MVPA, and thus will underrepresent the youths' achievements and awareness of their true PALs and goal attainment (ie, meeting the MVPA guideline bar), all of which could lead to negative feelings of the self [100]. To counteract such problems, future research should look to implement 3D-printed feedback with wrist-worn, fully waterproof accelerometers as they elicit higher wear-time compliance in youth than hip-mounted devices [133], as well as utilizing diary logs to account for contact sport activities [134]. Nevertheless, it is important to acknowledge that efforts are currently being made to make 3D printers faster, more accurate, and cheaper [135], with the potential for future research to involve youth more in the 3D printing process. Indeed, the number of schools owning a 3D printer is on the rise [136], which makes 3D-printing interventions similar to this study more feasible and cost-effective. In this light, it may be useful to compare 3D-printed feedback with other approaches, such as digital mobile phone feedback [137,138], light-emitting diode feedback technology [139,140], 3D-printed edibles [141], or shape-changing artifacts [142] to determine which methods of feedback can elicit the best intervention effects, user experience, and cost-effectiveness.

According to Forlizzi and Battarbee [47], new research methods are required to better articulate the relationship between what *we feel* and what *we do* in connection to the utilization of technology. This study builds on this by illustrating a short video interview approach to eliciting how youth experienced the 3D-printed models internally, functionally, and socially, all of which is essential for the development and future utilization of the designed 3D models [48]. The short video interviews generated a large set of descriptive data that could be generalized to the study population or used to account for an individual's personal progress and experiences with the 3D models, which aligns with the current trend toward *personalization* in health care [143] and the *quantified-self* movement [144]. However, one possible limitation to this aforementioned approach could be the direct influence of the ongoing short video interviews on the youths' experiences with the 3D models, given that previous research suggests that face-to-face support can create a more meaningful experience by reinforcing effort and goals [145,146]. In this regard, it could be argued that the ongoing face-to-face short video interviews may have potentially influenced youth awareness and motivation for PA, rather than the 3D models per se. Indeed, there are a number of practical ways a researcher or health professional could be deployed to support such a feedback intervention; however, to make technology-based behavior change strategies more pragmatic and cost-effective, it would be useful to understand the efficacy of support through continuous interviews [147]. Therefore, future research should look to break down 3D-printed feedback conditions to include and exclude face-to-face interviews to fully understand the impact of the tangible feedback and the influence of regular interviews on outcome variables [148]. That said, this study supports the use of short video interviews as a practical method for assessing the youths' experiences, understanding of, and interactions with the newly designed technology.

XSL•FO
RenderX

Limitations

There are, however, some additional limitations to consider to the aforementioned, such as the localized area of data collection in South Wales, which may underrepresent the ideologies of youth from other important socioeconomic groups and ethnic minorities in the United Kingdom or at a global level. Given the paucity of research on 3D-printed feedback, further research is required that considers the influence of age and sex, specifically, which may be hypothesized to influence initial engagement with the models. Indeed, the lack of a control group within this study questions whether the changes observed can be attributed to the impact of the 3D models per se to enhance youth awareness, goal setting, and motivation, and therefore, findings should be considered with caution and act as a stimulus for future investigation. Finally, this study was only a 7-week intervention with no long-term follow-up; therefore, it is unknown to what extent the novelty effect of the 3D models may diminish with time, as previously observed in the youths with wearable activity trackers [92]; therefore, a long-term study is warranted.

Conclusions

In conclusion, this study demonstrated that the age-specific 3D models heightened youth awareness of their PALs and enabled them to easily compare their personal PALs with the recommended guideline of 60 min of MVPA. Moreover, the youths expressed how they displayed their 3D models in their environments, within their bedrooms or next to prized possessions, and utilized the model as a goal-setting strategy to do more PA. Therefore, the nature of the age-specific 3D models being a combination of feedback, reward, and goal attainment that embodies personalized data may offer a unique strategy for the promotion of PA and associations to the recommended government guideline.

Acknowledgments

The authors thank all the teachers for their help and the children who participated in the study. The authors would also like to express their gratitude to Dr Ian Cameron and Andrew Fellows of Swansea University for their 3D printing knowledge and support.

Conflicts of Interest

None declared.

References

1. Department of Health, Physical Activity, Health Improvement and Protection. Gov.UK. 2011. Start Active, Stay Active: a report on physical activity for health from the four home countries' Chief Medical Officers URL: https://www.gov.uk/government/publications/start-active-stay-active-a-report-on-physical-activity-from-the-four-home-countries-chief-medical-officers [accessed 2019-05-13] [WebCite Cache ID 78LycHDoO]
2. Sothern M, Loftin M, Suskind RM, Udall JN, Blecker U. The health benefits of physical activity in children and adolescents: implications for chronic disease prevention. Eur J Pediatr 1999 Apr;158(4):271-274. [Medline: 10206121]
3. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act 2010 May 11;7:40 [FREE Full text] [doi: 10.1186/1479-5868-7-40] [Medline: 20459784]
4. Nieman P. Psychosocial aspects of physical activity. Paediatr Child Health 2002 May;7(5):309-312 [FREE Full text] [Medline: 20046307]
5. Eime R, Young JA, Harvey JT, Charity MJ, Payne WR. A systematic review of the psychological and social benefits of participation in sport for children and adolescents: informing development of a conceptual model of health through sport. Int J Behav Nutr Phys Act 2013 Aug 15;10:98 [FREE Full text] [doi: 10.1186/1479-5868-10-98] [Medline: 23945179]
6. NHS Digital. 2016. Health Survey for England: Children's health URL: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2016 [accessed 2019-05-13] [WebCite Cache ID 78Lyx9eiv]
7. NHS Digital. 2009. Health Survey for England - 2008: Physical activity and fitness URL: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2008-physical-activity-and-fitness [accessed 2019-05-13] [WebCite Cache ID 78Lz9oI0L]
8. Kremers S, Dijkman MA, de Meij JS, Jurg ME, Brug J. Awareness and habit: important factors in physical activity in children. Health Educ 2008 Oct 17;108(6):475-488. [doi: 10.1108/09654280810910881]
9. Corder K, van Sluijs EM, McMinn AM, Ekelund U, Cassidy A, Griffin SJ. Perception versus reality awareness of physical activity levels of British children. Am J Prev Med 2010 Jan;38(1):1-8 [FREE Full text] [doi: 10.1016/j.amepre.2009.08.025] [Medline: 20117551]
10. Xu F, Wang X, Xiang D, Wang Z, Ye Q, Ware RS. Awareness of knowledge and practice regarding physical activity: a population-based prospective, observational study among students in Nanjing, China. PLoS One 2017;12(6):e0179518 [FREE Full text] [doi: 10.1371/journal.pone.0179518] [Medline: 28622354]

XSL•FO
RenderX

11. Knox E, Esliger DW, Biddle SJ, Sherar LB. Lack of knowledge of physical activity guidelines: can physical activity promotion campaigns do better? BMJ Open 2013 Dec 5;3(12):e003633 [FREE Full text] [doi: 10.1136/bmjopen-2013-003633] [Medline: 24319271]
12. Snethen J, Broome ME. Weight, exercise, and health: children's perceptions. Clin Nurs Res 2007 May;16(2):138-152. [doi: 10.1177/1054773806298508] [Medline: 17452432]
13. Pearce PF, Harrell JS, McMurray RG. Middle-school children's understanding of physical activity:. J Pediatr Nurs 2008 Jun;23(3):169-182. [doi: 10.1016/j.pedn.2007.09.003] [Medline: 18492546]
14. Noonan R, Boddy LM, Fairclough SJ, Knowles ZR. Write, draw, show, and tell: a child-centred dual methodology to explore perceptions of out-of-school physical activity. BMC Public Health 2016 Apr 14;16:326 [FREE Full text] [doi: 10.1186/s12889-016-3005-1] [Medline: 27080384]
15. Cowden R, Plowman SA. The self-regulation and perception of exercise intensity in children in a field setting. Pediatr Exerc Sci 1999 Feb;11(1):32-43. [doi: 10.1123/pes.11.1.32]
16. Prochaska J, Sallis JF, Long B. A physical activity screening measure for use with adolescents in primary care. Arch Pediatr Adolesc Med 2001 May;155(5):554-559. [Medline: 11343497]
17. Placek J, Griffin LL, Dodds P, Raymond C, Tremino F, James A. Chapter 3: Middle school students' conceptions of fitness: the long road to a healthy lifestyle. J Teach Phys Educ 2001 Jul;20(4):314-323. [doi: 10.1123/jtpe.20.4.314]
18. Harris J, Cale L, Duncombe R, Musson H. Young people's knowledge and understanding of health, fitness and physical activity: issues, divides and dilemmas. Sport Educ Soc 2016 Sep 3;23(5):407-420. [doi: 10.1080/13573322.2016.1228047]
19. Godino J, Watkinson C, Corder K, Sutton S, Griffin SJ, van Sluijs EM. Awareness of physical activity in healthy middle-aged adults: a cross-sectional study of associations with sociodemographic, biological, behavioural, and psychological factors. BMC Public Health 2014 May 2;14:421 [FREE Full text] [doi: 10.1186/1471-2458-14-421] [Medline: 24886612]
20. Telama R. Tracking of physical activity from childhood to adulthood: a review. Obes Facts 2009;2(3):187-195 [FREE Full text] [doi: 10.1159/000222244] [Medline: 20054224]
21. Telama R, Yang X, Viikari J, Välimäki I, Wanne O, Raitakari O. Physical activity from childhood to adulthood: a 21-year tracking study. Am J Prev Med 2005 Apr;28(3):267-273. [doi: 10.1016/j.amepre.2004.12.003] [Medline: 15766614]
22. Weinstein N. The precaution adoption process. Health Psychol 1988;7(4):355-386. [Medline: 3049068]
23. Prochaska J, DiClemente CC. Stages of change in the modification of problem behaviors. Prog Behav Modif 1992;28:183-218. [Medline: 1620663]
24. Locke E, Latham GP. A Theory of Goal Setting & Task Performance. New Jersey, United States: Prentice-Hall; 1990:544.
25. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
26. Van Hoye K, Boen F, Lefevre J. The effects of physical activity feedback on behavior and awareness in employees: study protocol for a randomized controlled trial. Int J Telemed Appl 2012;2012:460712 [FREE Full text] [doi: 10.1155/2012/460712] [Medline: 23056040]
27. Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. Br Med J 2002 Apr 6;324(7341):827-830 [FREE Full text] [doi: 10.1136/bmj.324.7341.827] [Medline: 11934777]
28. Houts P, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006 May;61(2):173-190. [doi: 10.1016/j.pec.2005.05.004] [Medline: 16122896]
29. Bloch M. Truth and sight: generalizing without universalizing. J Royal Anthropological Inst 2008 Apr;14(s1):S22-S32. [doi: 10.1111/j.1467-9655.2008.00490.x]
30. Viseu A, Suchman L. Wearable augmentations: imaginaries of the informed body. In: Technologized Images, Technologized Bodies. New York: Berghahn Books; 2010:161-184.
31. Kreuter MW, Farrell DW, Brennan LK, Olevitch LR. Tailoring health messages: customizing communication with computer technology. In: Psycho-Oncology. New York: Routledge; 2013.
32. Ridgers N, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129 [FREE Full text] [doi: 10.2196/mhealth.6540] [Medline: 27881359]
33. Dean P, Voss C, De Souza A, Harris K. Assessment, Assurance, and Actuation: the use of activity trackers to monitor physical activity in a pediatric population with congenital heart disease. Can J Cardiol 2016 Oct;32(10):S130-S131. [doi: 10.1016/j.cjca.2016.07.194]
34. Hayes L, van Camp CM. Increasing physical activity of children during school recess. J Appl Behav Anal 2015 Sep;48(3):690-695. [doi: 10.1002/jaba.222] [Medline: 26119136]
35. Jacobsen R, Ginde S, Mussatto K, Neubauer J, Earing M, Danduran M. Can a home-based cardiac physical activity program improve the physical function quality of life in children with fontan circulation? Congenit Heart Dis 2016;11(2):175-182. [doi: 10.1111/chd.12330] [Medline: 26879633]
36. Schaefer S, Ching C, Breen H, German J. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. Am J Health Educ 2016 Jan 8;47(1):8-16. [doi: 10.1080/19325037.2015.1111174]

XSL•FO
RenderX

37. Hooke M, Gilchrist L, Tanner L, Hart N, Withycombe JS. Use of a fitness tracker to promote physical activity in children with acute lymphoblastic leukemia. Pediatr Blood Cancer 2016 Apr;63(4):684-689. [doi: 10.1002/pbc.25860] [Medline: 26756736]
38. Gaudet J, Gallant F, Bélanger M. A bit of fit: minimalist intervention in adolescents based on a physical activity tracker. JMIR Mhealth Uhealth 2017 Jul 6;5(7):e92 [FREE Full text] [doi: 10.2196/mhealth.7647] [Medline: 28684384]
39. Khot R, Mueller FF, Hjorth L. SweatAtoms: materializing physical activity. 2013 Presented at: The 9th Australasian Conference on Interactive Entertainment; September 30-October 1, 2013; Melbourne. [doi: 10.1145/2513002.2513012]
40. Price S, Rogers Y, Scaife M, Stanton D, Neale H. Using 'tangibles' to promote novel forms of playful learning. Interact Comput 2003 Apr;15(2):169-185. [doi: 10.1016/S0953-5438(03)00006-7]
41. Rogers Y, Scaife M, Gabrielli S, Smith H, Harris E. A conceptual framework for mixed reality environments: designing novel learning activities for young children. Presence (Camb) 2002 Dec;11(6):677-686. [doi: 10.1162/105474602321050776]
42. Dunn RS, Dunn KJ. Learning styles/teaching styles: should they ... can they ... be matched? Educ Leadersh 1979:238-244 [FREE Full text]
43. Piaget J, Cook M. The Origins of Intelligence in Children Volume 8. New York: International Universities Press; 1952:438.
44. Fleming N, Mills C. Not another inventory, rather a catalyst for reflection. Improve Acad 2017 Oct 10;11(1):137-155. [doi: 10.1002/j.2334-4822.1992.tb00213.x]
45. Crossley S, McNarry MA, Hudson J, Eslambolchilar P, Knowles Z, Mackintosh KA. Perceptions of visualising physical activity as a 3D-printed object: formative study. J Med Internet Res 2019 Jan 30;21(1):e12064 [FREE Full text] [doi: 10.2196/12064] [Medline: 30698532]
46. Crossley S, McNarry MA, Rosenberg M, Knowles ZR, Eslambolchilar P, Mackintosh KA. Understanding youths' ability to interpret 3D-printed physical activity data and identify associated intensity levels: mixed-methods study. J Med Internet Res 2019 Feb 22;21(2):e11253 [FREE Full text] [doi: 10.2196/11253] [Medline: 30794204]
47. Forlizzi J, Battarbee K. Understanding experience in interactive systems. 2004 Presented at: 5th conference on Designing interactive systems; August 1-4, 2004; Cambridge, USA p. 261-268. [doi: 10.1145/1013115.1013152]
48. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Inf Technol 2006 Mar;25(2):91-97. [doi: 10.1080/01449290500330331]
49. Lopes AG. Using research methods in human computer interaction to design technology for resilience. J Inf Syst Technol Manage 2016;13(3):363-388. [doi: 10.4301/25S1807-17752016000300001]
50. Masteller B, Sirard J, Freedson P. The Physical Activity Tracker Testing in Youth (P.A.T.T.Y.) study: content analysis and children's perceptions. JMIR Mhealth Uhealth 2017 Apr 28;5(4):e55 [FREE Full text] [doi: 10.2196/mhealth.6347] [Medline: 28455278]
51. Marshall P. Do tangible interfaces enhance learning? In: Proceedings of the 1st international conference on Tangible and embedded interaction. 2007 Presented at: TEI'07; February 15-17, 2007; Baton Rouge, Louisiana p. 163-170. [doi: 10.1145/1226969.1227004]
52. Rogers Y, Scaife M, Gabrielli S, Smith H, Harris E. A conceptual framework for mixed reality environments: Designing novel learning activities for young children. Presence Teleoperators and Virtual Environments httpsdoiorg10116210547402321050776, 2002. 11(6 DOI - 10. ACM 2002;11(6):677-686.
53. Goodman J, Wood RE. Faded versus increasing feedback, task variability trajectories, and transfer of training. Hum Perform 2009 Jan 2;22(1):64-85. [doi: 10.1080/08959280802541013]
54. Day E, Blair C, Daniels S, Kligyte V, Mumford M. Linking instructional objectives to the design of instructional environments: the Integrative Training Design Matrix. Hum Res Manage Rev 2006 Sep;16(3):376-395. [doi: 10.1016/j.hrmr.2006.05.007]
55. Edwards L, Salant V, Howard VF, Brougher J, McLaughlin TF. Effectiveness of self-management on attentional behavior and reading comprehension for children with attention deficit disorder. Child Fam Behav Ther 1995 Sep 12;17(2):1-17. [doi: 10.1300/J019v17n02_01]
56. Hesketh B. Dilemmas in training for transfer and retention. Appl Psychol 1997;46(4):317-339. [doi: 10.1111/j.1464-0597.1997.tb01234]
57. Rock ML, Thead BK. The effects of fading a strategic self-monitoring intervention on students? academic engagement, accuracy, and productivity. J Behav Educ 2007;16(4):389-412 [FREE Full text]
58. Schmidt RA, Wrisberg CA. Motor Learning and Performance: A Situation-Based Learning Approach. USA: Human Kinetics; 2008.
59. 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 Apr 24;10:59 [FREE Full text] [doi: 10.1186/s13012-015-0238-9] [Medline: 25906822]
60. Kennedy C, Kools S, Krueger R. Methodological considerations in children's focus groups. Nurs Res 2001;50(3):184-187. [Medline: 11393641]
61. Stewart A, Marfell-Jones M, Olds T, de Ridder H. International Standards for Anthropometric Assessment. Glasgow, Scotland: International Society for the Advancement of Kinanthropometry; 2011:137.

XSL•FO
RenderX

62. Cole T, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. Br Med J 2000 May 6;320(7244):1240-1243 [FREE Full text] [doi: 10.1136/bmj.320.7244.1240] [Medline: 10797032]
63. The National Institute for Health and Care Excellence. 2014. Obesity: identification, assessment and management URL: https://www.nice.org.uk/guidance/cg189 [accessed 2019-05-10] [WebCite Cache ID 78GuIljVY]
64. Trost S. State of the art reviews: measurement of physical activity in children and adolescents. Am J Lifestyle Med 2016 Oct 12;1(4):299-314. [doi: 10.1177/1559827607301686]
65. Hills A, Mokhtar N, Byrne NM. Assessment of physical activity and energy expenditure: an overview of objective measures. Front Nutr 2014;1:5 [FREE Full text] [doi: 10.3389/fnut.2014.00005] [Medline: 25988109]
66. Tudor-Locke C, Barreira TV, Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. Med Sci Sports Exerc 2015 Apr;47(4):839-842. [doi: 10.1249/MSS.0000000000000476] [Medline: 25121517]
67. Cleland I, Kikhia B, Nugent C, Boytsov A, Hallberg J, Synnes K, et al. Optimal placement of accelerometers for the detection of everyday activities. Sensors (Basel) 2013 Jul 17;13(7):9183-9200 [FREE Full text] [doi: 10.3390/s130709183] [Medline: 23867744]
68. Robertson W, Stewart-Brown S, Wilcock E, Oldfield M, Thorogood M. Utility of accelerometers to measure physical activity in children attending an obesity treatment intervention. J Obes 2011;2011:8 [FREE Full text] [doi: 10.1155/2011/398918] [Medline: 20953356]
69. Evenson K, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci 2008 Dec;26(14):1557-1565. [doi: 10.1080/02640410802334196] [Medline: 18949660]
70. Crouter S, Horton M, Bassett DR. Validity of ActiGraph child-specific equations during various physical activities. Med Sci Sports Exerc 2013 Jul;45(7):1403-1409 [FREE Full text] [doi: 10.1249/MSS.0b013e318285f03b] [Medline: 23439413]
71. Catellier D, Hannan PJ, Murray DM, Addy CL, Conway TL, Yang S, et al. Imputation of missing data when measuring physical activity by accelerometry. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S555-S562 [FREE Full text] [Medline: 16294118]
72. Mattocks C, Ness A, Leary S, Tilling K, Blair SN, Shield J, et al. Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision. J Phys Act Health 2008;5(Suppl 1):S98-111. [Medline: 18364528]
73. Rich C, Geraci M, Griffiths L, Sera F, Dezateux C, Cortina-Borja M. Quality control methods in accelerometer data processing: defining minimum wear time. PLoS One 2013;8(6):e67206 [FREE Full text] [doi: 10.1371/journal.pone.0067206] [Medline: 23826236]
74. Morgan DL. Focus Groups as Qualitative Research. Thousand Oaks, CA: Sage publications; 1996.
75. Deanscombe M. The Good Research Guide for Small Scale Research Projects. Berkshire, UK: Open University Press; 2010.
76. Lewis A. Group child interviews as a research tool. Br Educ Res J 1992 Jan;18(4):413-421. [doi: 10.1080/0141192920180407]
77. Banks M. Using Visual Data in Qualitative Research. Oxford: Sage Publications; 2018.
78. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. Thousand Oaks, CA: Sage Publications; 2014:217.
79. Hamad E, Savundranayagam MY, Holmes JD, Kinsella EA, Johnson AM. Toward a mixed-methods research approach to content analysis in the digital age: the combined content-analysis model and its applications to health care Twitter feeds. J Med Internet Res 2016 Mar 8;18(3):e60 [FREE Full text] [doi: 10.2196/jmir.5391] [Medline: 26957477]
80. Clarke V, Braun V. Teaching thematic analysis: overcoming challenges and developing strategies for effective learning. Psychologist 2013;26(2):120-123 [FREE Full text]
81. Boddy L, Knowles ZR, Davies IG, Warburton GL, Mackintosh KA, Houghton L, et al. Using formative research to develop the healthy eating component of the CHANGE! school-based curriculum intervention. BMC Public Health 2012 Aug 29;12:710 [FREE Full text] [doi: 10.1186/1471-2458-12-710] [Medline: 22931457]
82. Mackintosh K, Knowles ZR, Ridgers ND, Fairclough SJ. Using formative research to develop CHANGE!: a curriculum-based physical activity promoting intervention. BMC Public Health 2011 Oct 27;11:831 [FREE Full text] [doi: 10.1186/1471-2458-11-831] [Medline: 22032540]
83. Winn C, Mackintosh KA, Eddolls WT, Stratton G, Wilson AM, Rance JY, et al. Perceptions of asthma and exercise in adolescents with and without asthma. J Asthma 2018 Aug;55(8):868-876. [doi: 10.1080/02770903.2017.1369992] [Medline: 28853952]
84. Vanhelst J, Béghin L, Drumez E, Coopman S, Gottrand F. Awareness of wearing an accelerometer does not affect physical activity in youth. BMC Med Res Methodol 2017 Jul 11;17(1):99 [FREE Full text] [doi: 10.1186/s12874-017-0378-5] [Medline: 28693500]
85. Bentley G, Goodred JK, Jago R, Sebire SJ, Lucas PJ, Fox KR, et al. Parents' views on child physical activity and their implications for physical activity parenting interventions: a qualitative study. BMC Pediatr 2012 Nov 20;12:180 [FREE Full text] [doi: 10.1186/1471-2431-12-180] [Medline: 23167910]
86. Health Survey for England. Digital NHS. 2007. 2007: Healthy lifestyles: knowledge, attitudes and behaviour URL: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/

XSL•FO
RenderX

health-survey-for-england-2007-healthy-lifestyles-knowledge-attitudes-and-behaviour [accessed 2019-05-14] [WebCite Cache ID 78MEPksmN]
87. Roth M, Stamatakis E. Linking young people's knowledge of public health guidelines to physical activity levels in England. Pediatr Exerc Sci 2010 Aug;22(3):467-476. [Medline: 20814041]
88. Nemet D, Geva D, Meckel Y, Eliakim A. Health-related knowledge and preferences in low socio-economic kindergarteners. Int J Behav Nutr Phys Act 2012 Jan 10;9:1 [FREE Full text] [doi: 10.1186/1479-5868-9-1] [Medline: 22233712]
89. Ronda G, van Assema P, Brug J. Stages of change, psychological factors and awareness of physical activity levels in The Netherlands. Health Promot Int 2001 Dec;16(4):305-314. [Medline: 11733449]
90. De Bourdeaudhuij I, Philippaerts R, Crombez G, Matton L, Wijndaele K, Balduck AL, et al. Stages of change for physical activity in a community sample of adolescents. Health Educ Res 2005 Jun;20(3):357-366. [doi: 10.1093/her/cyg131] [Medline: 15536126]
91. Golsteijn C, van den Hoven E, Frohlich D, Sellen A. Towards a more cherishable digital object. In: Proceedings of the Designing Interactive Systems Conference. 2012 Presented at: DIS'12; June 11-15, 2012; New York, USA p. 655-664.
92. Ridgers N, Timperio A, Brown H, Ball K, Macfarlane S, Lai SK, et al. Wearable activity tracker use among Australian adolescents: usability and acceptability study. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e86 [FREE Full text] [doi: 10.2196/mhealth.9199] [Medline: 29643054]
93. Maturo C, Cunningham SA. Influence of friends on children's physical activity: a review. Am J Public Health 2013 Jul;103(7):e23-e38 [FREE Full text] [doi: 10.2105/AJPH.2013.301366] [Medline: 23678914]
94. Pearce M, Page AS, Griffin TP, Cooper AR. Who children spend time with after school: associations with objectively recorded indoor and outdoor physical activity. Int J Behav Nutr Phys Act 2014 Mar 30;11(1):45 [FREE Full text] [doi: 10.1186/1479-5868-11-45] [Medline: 24679149]
95. Salvy S, Roemmich JN, Bowker JC, Romero ND, Stadler PJ, Epstein LH. Effect of peers and friends on youth physical activity and motivation to be physically active. J Pediatr Psychol 2009 Mar;34(2):217-225 [FREE Full text] [doi: 10.1093/jpepsy/jsn071] [Medline: 18617572]
96. Davison K, Mâsse LC, Timperio A, Frenn MD, Saunders J, Mendoza JA, et al. Physical activity parenting measurement and research: challenges, explanations, and solutions. Child Obes 2013 Aug;9(Suppl):S103-S109 [FREE Full text] [doi: 10.1089/chi.2013.0037] [Medline: 23944918]
97. Haerens L, De Bourdeaudhuij I, Maes L, Cardon G, Deforche B. School-based randomized controlled trial of a physical activity intervention among adolescents. J Adolesc Health 2007 Mar;40(3):258-265. [doi: 10.1016/j.jadohealth.2006.09.028] [Medline: 17321427]
98. Haerens L, Deforche B, Maes L, Cardon G, Stevens V, De Bourdeaudhuij I. Evaluation of a 2-year physical activity and healthy eating intervention in middle school children. Health Educ Res 2006 Dec;21(6):911-921. [doi: 10.1093/her/cyl115] [Medline: 17032704]
99. Sleddens S, Gerards SM, Thijs C, de Vries NK, Kremers SP. General parenting, childhood overweight and obesity-inducing behaviors: a review. Int J Pediatr Obes 2011 Jun;6(2-2):e12-e27. [doi: 10.3109/17477166.2011.566339] [Medline: 21657834]
100. Goodyear V, Kerner C, Quennerstedt M. Young people's uses of wearable healthy lifestyle technologies; surveillance, self-surveillance and resistance. Sport Educ Soc 2017 Sep 22;24(3):212-225. [doi: 10.1080/13573322.2017.1375907]
101. Harter S. Causes and consequences of low self-esteem in children and adolescents. In: Baumeister RF, editor. Self-Esteem. New York: Plenum Press; 1993:87-116.
102. Kerner C, Goodyear VA. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on Fitbits with adolescents. Am J Health Educ 2017 Jul 28;48(5):287-297. [doi: 10.1080/19325037.2017.1343161]
103. Khot RA. Understanding Material Representations of Physical Activity. Melbourne, Australia: RMIT University; 2016:144.
104. Cullen K, Baranowski T, Smith SP. Using goal setting as a strategy for dietary behavior change. J Am Diet Assoc 2001 May;101(5):562-566. [doi: 10.1016/S0002-8223(01)00140-7] [Medline: 11374350]
105. Locke E, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [Medline: 12237980]
106. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Decis Process 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-L]
107. Hynynen S, van Stralen MM, Sniehotta FF, Araújo-Soares V, Hardeman W, Chinapaw MJ, et al. A systematic review of school-based interventions targeting physical activity and sedentary behaviour among older adolescents. Int Rev Sport Exerc Psychol 2016 Jan 1;9(1):22-44 [FREE Full text] [doi: 10.1080/1750984X.2015.1081706] [Medline: 26807143]
108. Kyllo LB, Landers DM. Goal setting in sport and exercise: a research synthesis to resolve the controversy. J Sport Exerc Psychol 1995 Jun;17(2):117-137. [doi: 10.1123/jsep.17.2.117]
109. Rose T, Barker M, Maria Jacob C, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]
110. Weinberg R, Weigand D. Goal setting in sport and exercise: a reaction to Locke. J Sport Exerc Psychol 1993 Mar;15(1):88-96. [doi: 10.1123/jsep.15.1.88]

XSL•FO
RenderX

111. Ivancevich J, McMahon J. The effects of goal setting, external feedback, and self-generated feedback on outcome variables: a field experiment. AMJ 1982 Jun;25(2):359-372. [doi: 10.5465/255997]
112. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr 27;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
113. Latham G, Locke EA. Self-regulation through goal setting. Organ Behav Hum Decis Process 1991 Dec;50(2):212-247. [doi: 10.1016/0749-5978(91)90021-K]
114. Locke E, Latham GP. New directions in goal-setting theory. Curr Dir Psychol Sci 2016 Jun 23;15(5):265-268. [doi: 10.1111/j.1467-8721.2006.00449.x]
115. Hardman C, Horne PJ, Fergus Lowe C. Effects of rewards, peer-modelling and pedometer targets on children's physical activity: a school-based intervention study. Psychol Health 2011 Jan;26(1):3-21. [doi: 10.1080/08870440903318119] [Medline: 20204975]
116. Finkelstein E, Tan YT, Malhotra R, Lee CF, Goh SS, Saw SM. A cluster randomized controlled trial of an incentive-based outdoor physical activity program. J Pediatr 2013 Jul;163(1):167-72.e1. [doi: 10.1016/j.jpeds.2013.01.009] [Medline: 23415616]
117. Christian D, Todd C, Hill R, Rance J, Mackintosh K, Stratton G, et al. Active children through incentive vouchers - evaluation (ACTIVE): a mixed-method feasibility study. BMC Public Health 2016 Dec 26;16(1):890 [FREE Full text] [doi: 10.1186/s12889-016-3381-6] [Medline: 27566535]
118. Lepper M, Greene D, Nisbett RE. Undermining children's intrinsic interest with extrinsic reward: A test of the. J Pers Soc Psychol 1973;28(1):129-137. [doi: 10.1037/h0035519]
119. Ryan R, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
120. Corepal R, Tully MA, Kee F, Miller SJ, Hunter RF. Behavioural incentive interventions for health behaviour change in young people (5-18 years old): a systematic review and meta-analysis. Prev Med 2018 Dec;110:55-66. [doi: 10.1016/j.ypmed.2018.02.004] [Medline: 29432789]
121. Shilts M, Horowitz M, Townsend MS. Goal setting as a strategy for dietary and physical activity behavior change: a review of the literature. Am J Health Promot 2004;19(2):81-93. [doi: 10.4278/0890-1171-19.2.81] [Medline: 15559708]
122. Cole M, Wertsch JV. Beyond the individual-social antinomy in discussions of Piaget and Vygotsky. Hum Dev 1996;39(5):250-256. [doi: 10.1159/000278475]
123. Baranowski T. Validity and reliability of self report measures of physical activity: an information-processing perspective. Res Q Exerc Sport 1988 Dec;59(4):314-327. [doi: 10.1080/02701367.1988.10609379]
124. Sallis J. Self-report measures of children's physical activity. J Sch Health 1991 May;61(5):215-219. [Medline: 1943046]
125. Husmann P, O'Loughlin VD. Another nail in the coffin for learning styles? Disparities among undergraduate anatomy students? study strategies, class performance, and reported VARK learning styleS. Anat Sci Educ 2019 Jan;12(1):6-19. [doi: 10.1002/ase.1777] [Medline: 29533532]
126. Coffield F, Moseley D, Hall E, Ecclestone K. Should we be using learning styles? What research has to say to practice. Learn Skills Res Centre 2004:84 [FREE Full text]
127. Papanagnou D, Serrano A, Barkley K, Chandra S, Governatori N, Piela N, et al. Does tailoring instructional style to a medical student's self-perceived learning style improve performance when teaching intravenous catheter placement? A randomized controlled study. BMC Med Educ 2016 Aug 12;16(1):205 [FREE Full text] [doi: 10.1186/s12909-016-0720-3] [Medline: 27520578]
128. Pashler H, McDaniel M, Rohrer D, Bjork R. Learning styles: concepts and evidence. Psychol Sci Public Interest 2008 Dec;9(3):105-119. [doi: 10.1111/j.1539-6053.2009.01038.x] [Medline: 26162104]
129. Riener C, Willingham D. The myth of learning styles. Change Magazine Higher Learn 2010 Aug 30;42(5):32-35. [doi: 10.1080/00091383.2010.503139]
130. Biddle S, Gorely T, Marshall SJ, Cameron N. The prevalence of sedentary behavior and physical activity in leisure time: a study of Scottish adolescents using ecological momentary assessment. Prev Med 2009 Feb;48(2):151-155. [doi: 10.1016/j.ypmed.2008.10.025] [Medline: 19046984]
131. Corder K, Ekelund U, Steele RM, Wareham NJ, Brage S. Assessment of physical activity in youth. J Appl Physiol (1985) 2008 Sep;105(3):977-987 [FREE Full text] [doi: 10.1152/japplphysiol.00094.2008] [Medline: 18635884]
132. Gorely T, Biddle SJ, Marshall SJ, Cameron N. The prevalence of leisure time sedentary behaviour and physical activity in adolescent boys: an ecological momentary assessment approach. Int J Pediatr Obes 2009;4(4):289-298. [doi: 10.3109/17477160902811181] [Medline: 19922044]
133. Trost S, Zheng Y, Wong WK. Machine learning for activity recognition: hip versus wrist data. Physiol Meas 2014 Nov;35(11):2183-2189. [doi: 10.1088/0967-3334/35/11/2183] [Medline: 25340887]
134. Pfitzner R, Gorzelniak L, Heinrich J, von Berg A, Klümper C, Bauer CP, GINIplus Study Group. Physical activity in German adolescents measured by accelerometry and activity diary: introducing a comprehensive approach for data management and preliminary results. PLoS One 2013;8(6):e65192 [FREE Full text] [doi: 10.1371/journal.pone.0065192] [Medline: 23750243]

XSL•FO
RenderX

135. Mueller S, Mohr T, Guenther K, Frohnhofen J, Baudisch P. faBrickation: fast 3D printing of functional objects by integrating construction kit building blocks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI EA'14; April 26-May 1, 2014; Toronto, Canada.
136. Department of Education. Digital Education Resource Archive. 2013. 3D printers in schools: uses in the curriculum: enriching the teaching of STEM and design subjects URL: https://dera.ioe.ac.uk/19468/https://dera.ioe.ac.uk/19468/ [accessed 2019-05-14] [WebCite Cache ID 78MooTYZo]
137. Larsen JE, Cuttone A, Jørgensen SL. QS Spiral: Visualizing Periodic Quantified Self Data. In: Proceedings of CHI 2013 Workshop on Personal Informatics in the Wild: Hacking Habits for Health & Happiness. 2013 Presented at: CHI 2013; 27 April-2 May, 2013; Paris, France.
138. Tong X, Gromala D, Bartman L, Rajabiyazdi F, Carpendale S. Evaluating the Effectiveness of Three Physical Activity Visualizations — How People Perform vs. Perceive. 2015 Presented at: IEEE VIS Personal Visualization Workshop; October 25, 2015; Chicago, Illinois, USA p. 4.
139. Ananthanarayan S, Siek K, Eisenberg M. A Craft Approach to Health Awareness in Children. In: Proceedings of the 2016 ACM Conference on Designing Interactive Systems. 2016 Presented at: DIS'16; June 4-8, 2016; Brisbane, Australia p. 724-735. [doi: 10.1145/2901790.2901888]
140. Mackintosh K, Niezen G, Eslambolchilar P. Mission possible: using ubiquitous social goal sharing technology to promote physical activity in children. Malaysian J Move Health Exerc 2016 Oct 18;5(2):- [FREE Full text] [doi: 10.15282/mohe.v5i2.115]
141. Rohit AK, Ryan P, Mueller FF. EdiPulse: Turning Physical Activity Into Chocolates. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. 2015 Presented at: CHI EA'15; April 18-23, 2015; Seoul, Republic of Korea.
142. Sauvé K, Houben S, Marquardt N, Bakker S, Hengeveld B, Gallacher S, et al. LOOP: A Physical Artifact to Facilitate Seamless Interaction with Personal Data in Everyday Life. In: Proceedings of the ACM Conference Companion Publication on Designing Interactive Systems. 2017 Presented at: DIS'17 Companion; June 1014, 2017; Edinburgh, UK.
143. Cesuroglu T, Syurina E, Feron F, Krumeich A. Other side of the coin for personalised medicine and healthcare: content analysis of 'personalised' practices in the literature. BMJ Open 2016 Dec 13;6(7):e010243 [FREE Full text] [doi: 10.1136/bmjopen-2015-010243] [Medline: 27412099]
144. Almalki M, Gray K, Sanchez FM. The use of self-quantification systems for personal health information: big data management activities and prospects. Health Inf Sci Syst 2015;3(Suppl 1 HISA Big Data in Biomedicine and Healthcare 2013 Con):S1 [FREE Full text] [doi: 10.1186/2047-2501-3-S1-S1] [Medline: 26019809]
145. Fortier M, Duda JL, Guerin E, Teixeira PJ. Promoting physical activity: development and testing of self-determination theory-based interventions. Int J Behav Nutr Phys Act 2012 Mar 2;9(1):20 [FREE Full text] [doi: 10.1186/1479-5868-9-20] [Medline: 22385751]
146. Williams S, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? Health Educ Res 2011 Apr;26(2):308-322. [doi: 10.1093/her/cyr005] [Medline: 21321008]
147. Heath G, Parra DC, Sarmiento OL, Andersen LB, Owen N, Goenka S, Lancet Physical Activity Series Working Group. Evidence-based intervention in physical activity: lessons from around the world. Lancet 2012 Jul 21;380(9838):272-281 [FREE Full text] [doi: 10.1016/S0140-6736(12)60816-2] [Medline: 22818939]
148. Moore G, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]

Abbreviations

3D: three-dimensional
BMI: body mass index
M1: model 1
M2: model 2
M3: model 3
M4: model 4
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PALs: physical activity levels
PAPM: Precaution Adoption Process Model

XSL•FO
RenderX

Edited by G Eysenbach; submitted 29.08.18; peer-reviewed by A Chwałczyńska, Y Wang, S Jalil, H Oh; comments to author 02.02.19; revised version received 13.03.19; accepted 31.03.19; published 31.05.19.

Please cite as:
Crossley SGM, McNarry MA, Eslambolchilar P, Knowles Z, Mackintosh KA
The Tangibility of Personalized 3D-Printed Feedback May Enhance Youths' Physical Activity Awareness, Goal Setting, and Motivation: Intervention Study
J Med Internet Res 2019;21(6):e12067
URL: https://www.jmir.org/2019/6/e12067/
doi:10.2196/12067
PMID:31199322

©Sam Graeme Morgan Crossley, Melitta Anne McNarry, Parisa Eslambolchilar, Zoe Knowles, Kelly Alexandra Mackintosh. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 31.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Accelerating Health Data Sharing: A Solution Based on the Internet of Things and Distributed Ledger Technologies

Xiaochen Zheng[1], DrPH; Shengjing Sun[1], MSci; Raghava Rao Mukkamala[2,3], DPhil; Ravi Vatrapu[2,3], DPhil; Joaquín Ordieres-Meré[1], DPhil

[1]Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, Madrid, Spain

[2]Centre for Business Data Analytics, Department of Digitalization, Copenhagen Business School, Copenhagen, Denmark

[3]Department of Technology, Kristiania University College, Oslo, Norway

Corresponding Author:
Joaquín Ordieres-Meré, DPhil
Escuela Técnica Superior de Ingenieros Industriales
Universidad Politécnica de Madrid
José Gutiérrez Abascal 2
Madrid, 28006
Spain
Phone: 34 910677107
Fax: 34 913363005
Email: j.ordieres@upm.es

Abstract

Background: Huge amounts of health-related data are generated every moment with the rapid development of Internet of Things (IoT) and wearable technologies. These big health data contain great value and can bring benefit to all stakeholders in the health care ecosystem. Currently, most of these data are siloed and fragmented in different health care systems or public and private databases. It prevents the fulfillment of intelligent health care inspired by these big data. Security and privacy concerns and the lack of ensured authenticity trails of data bring even more obstacles to health data sharing. With a decentralized and consensus-driven nature, distributed ledger technologies (DLTs) provide reliable solutions such as blockchain, Ethereum, and IOTA Tangle to facilitate the health care data sharing.

Objective: This study aimed to develop a health-related data sharing system by integrating IoT and DLT to enable secure, fee-less, tamper-resistant, highly-scalable, and granularly-controllable health data exchange, as well as build a prototype and conduct experiments to verify the feasibility of the proposed solution.

Methods: The health-related data are generated by 2 types of IoT devices: wearable devices and stationary air quality sensors. The data sharing mechanism is enabled by IOTA's distributed ledger, the Tangle, which is a directed acyclic graph. Masked Authenticated Messaging (MAM) is adopted to facilitate data communications among different parties. Merkle Hash Tree is used for data encryption and verification.

Results: A prototype system was built according to the proposed solution. It uses a smartwatch and multiple air sensors as the sensing layer; a smartphone and a single-board computer (Raspberry Pi) as the gateway; and a local server for data publishing. The prototype was applied to the remote diagnosis of tremor disease. The results proved that the solution could enable costless data integrity and flexible access management during data sharing.

Conclusions: DLT integrated with IoT technologies could greatly improve the health-related data sharing. The proposed solution based on IOTA Tangle and MAM could overcome many challenges faced by other traditional blockchain-based solutions in terms of cost, efficiency, scalability, and flexibility in data access management. This study also showed the possibility of fully decentralized health data sharing by replacing the local server with edge computing devices.

(J Med Internet Res 2019;21(6):e13583) doi:10.2196/13583

KEYWORDS

Internet of Things; distributed ledger technologies; data sharing; health information interoperability; IOTA Tangle; masked authenticated messaging; blockchain; intelligent healthcare

XSL•FO
RenderX

Introduction

Internet of Things and Intelligent Health Care

Internet of Things (IoT) has been developing explosively in recent years. It is believed to be the next revolutionary technology and bring great benefits to various domains of the society including health care [1]. The health care industry has been dramatically changed because of the information technology revolution that started in the last century. New technologies such as telemedicine, digital hospital, electronic health and mobile health have been widely applied during the past decades, and now, the rapidly development of IoT is promoting health care from digital into intelligent [2].

The advances of IoT have resulted in rapid emergence of smart environments such as smart home [3]. Sensors in these environments can measure the values of various environmental factors including temperature, humidity, air quality, and noise [4].

As an important aspect of IoT, wearable technology has also shown a surge in the past decade. Different types of wearable devices containing various embedded sensors such as smartphone, smart watch, smart band, and smart glasses have been used in health care applications to realize various health-related applications such as remote diagnosis [5], disease monitoring [6], and elderly people caring [7].

Challenges of Health Care Data Sharing

Large amounts of health-related data are generated by these smart devices including environmental data from stationary sensors and activity data from wearable devices. These data are valuable resources for health care applications, research, and commercial projects. Properly sharing these health data can benefit all related stakeholders including the device users, patients, researchers, and companies and improve the public health care system.

Currently, most data generated by IoT devices are controlled by different service providers, device manufacturers, or scattered in different health care systems [8,9]. These siloed and segmented data make it impossible or very difficult to share data outside their own closed environments, and this leads to enormous quantities of wasted data [10]. Besides, it puts data security and privacy at risk as these centralized data stores and authority providers are attractive targets for cyberattacks [11].

With the increasing concern about data privacy and security issue from public and private users, data protection regulations will become stricter. For example, the European Union has published the General Data Protection Regulation [12] to protect individual data. Such regulations make data sharing even more difficult.

Besides the complex data protection regulations, another main obstacle to freely flowing of big data is that, although data sharing is becoming cheaper from a technological perspective, it is prohibitively expensive to transfer fine, granular data in real time because of intermediary fees [13]. Another barrier is the lack of ensured authenticity and audit trails of data. Traditional data transmission protocols and databases are susceptible to various attacks, including *man-in-the-middle* attacks and data tampering [14].

To overcome these barriers that hinder the full use of valuable health data, it is necessary to develop advanced systems to accelerate secure, fee-less, tamper-resist, and high-scalable health data sharing.

Distributed Ledger Technologies and Blockchain

A distributed ledger is a distributed database, maintained by a consensus protocol run by nodes in a peer-to-peer network. This consensus protocol replaces a central administrator, as all peers contribute to maintaining the integrity of the database [15].

As one of the most widespread DLT, the blockchain, has gained substantial popularity in recent years, primarily in the financial field because of the cryptocurrencies. For example, Bitcoin was first introduced in 2008 [16] and ever since has attracted the attention of the research community from diverse academic fields [17-19] and gained mainstream popularity because of its unique characteristics such as the absence of centralized control, an assumed high degree of anonymity, and distributed consensus over decentralized networks. Blockchain solutions could reduce data breach risks by utilizing threshold encryption of data together using public key infrastructure, where cooperation of multiple parties is required to decrypt data, and asymmetric cryptography is used to authenticate communication with system participants [20].

Limitations of Blockchain

Specialized distributed consensus protocols based on DLT have enabled novel decentralized applications such as cryptographic currencies [16] and smart contracts [21]. The rise and success of Bitcoin during the last 6 years proved that blockchain technology has real-world value. However, these block-based protocols, such as blockchain and Ethereum, also have several drawbacks that prevent them from being used as a generic platform for IoT data sharing.

Scalability

A blockchain has an inherent transaction rate limit because all participants agree on the longest chain and discard forks and side branches [22]. Common practice is to wait for 6 blocks to be added to the longest chain before reaching a high level of confidence that a transaction is final on the Bitcoin network [20,23]. As an example, it took on average 9.3 min to confirm a Bitcoin transaction at the end of December 2018 [24]. Applications that require exchange of value and low latency cannot be certain that their transactions are final in a shorter time frame and must trust the payer to not double spend [20]. The current incentive schemes that allow these protocols to spread virally make inefficient use of computational resources while constraining the transaction rate on the network. The transaction rate of Bitcoin protocol has been lower than 5 transactions per second in the whole network during most of the time in the year 2018 [25]. Similarly, the Ethereum protocol currently processes about 6 transactions per second across the entire network [26]. This low throughput cannot fill the requirements of data sharing in many health care scenarios.

XSL•FO
RenderX

Fees

Another notable drawback is the concept of a transaction fee for transactions of any value. For example, the Bitcoin protocol requires a fee that may exceed US $0.30 each transaction [27] according to the statistics of January 10, 2019. To use a distributed ledger at scale for financial or other industrial use cases, this low throughput and high fee model will not suffice. The importance of micropayments will increase in the rapidly developing IoT technology and paying a fee that is larger than the amount of value being transferred is not logical. Furthermore, it is not easy to get rid of fees in the blockchain infrastructure as they serve as an incentive for the creators of blocks [28].

Centralization

Lots of computing power is required to maintain the blockchain, and mining power has become centralized to some extent. The latest statistic shows that the 6 largest mining pools control 75.76% of the of the network's mining power (BTC.com 21.5%, AntPool 14.9%, SlushPool 11.01%, ViaBTC 10.65%, BTC.TOP 9.67%, F2Pool 8.03%) [29].

Vulnerable to Quantum Attack

Bitcoin and other proof-of-work–based blockchains are susceptible to being broken by quantum computers. Quantum computers, although still a hypothetical construct as of today, could be very efficient for handling problems that rely on trial and error to find a solution [28]. The process of finding a nonce to generate a Bitcoin block is a good example of such a problem. As of today, one must check an average of 2^{68} nonce to find a suitable hash that allows a new block to be generated. Theoretically, a quantum computer would need $\theta(\sqrt{N})$ operations to solve a problem that is analogous to the Bitcoin puzzle stated above [30]. This same problem would need $\theta(\sqrt{N})$ operations on a classical computer. Therefore, a quantum computer would be around 17 billion ($\sqrt{2^{68}}$) times more efficient at mining the Bitcoin blockchain than a classical computer. It would make possible of gaining control of over 51% of computing power of the whole blockchain network, which would enable attackers to double spend and break the entire network.

IOTA and the Tangle

IOTA is a tangle-based cryptocurrency designed specifically for the IoT industry where a machine-to-machine micropayment system is required. The tangle naturally succeeds the blockchain as its next evolutionary step by overcoming some of its previously mentioned fundamental limitations [31]. The main feature of the tangle is that it uses a directed acyclic graph for storing transactions instead of sequential blocks. In the Tangle, users must perform a small amount of computational work to approve 2 previous transactions to issue a new transaction. This new transaction will be validated by some subsequent transactions [28].

This structure enables the Tangle with high scalability. There is no maximum throughput, as the more activities in the Tangle, the faster transactions can be confirmed. In addition, with this 'pay-it-forward' system of validations, there is no need to offer financial rewards. Transacting with IOTA can be free of charge. Moreover, IOTA has no miners, therefore it is truly decentralized.

The IOTA tangle is designed to be quantum resistant. The number of nonce that one needs to check to find a suitable hash for issuing a transaction is around 3^8 on average, which is not unreasonably large. The gain of efficiency for an "ideal" quantum computer would therefore be of order 3^4=81, which is already quite acceptable [28]. More importantly, the algorithm used in the IOTA implementation is structured such that the time to find a nonce is not much longer than the time needed for other tasks that are necessary to issue a transaction. The latter part is much more resistant against quantum computing compared with the traditional blockchain.

Masked Authenticated Messaging

The main data communication protocol used in the proposed system is Masked Authenticated Messaging (MAM). It enables to emit and access an encrypted data stream over the Tangle regardless of the size or cost of a device [32]. MAM uses channels for message spreading. IOTA users can create a channel and publish a message of any size at any time. A small amount of proof-of-work is required to allow the data to propagate through the network and to prevent spamming. Other users can subscribe this channel through its address and receive a message that is published by the channel owner.

Merkle Hash Tree

MAM uses a signature scheme based on Merkle Hash Tree (MHT) [33-35] to sign the cipher digest of an encrypted message [32]. The *address* of a channel is the *root* of this Merkle tree, which itself is created using the *seed* of the user.

As the MHT example shown in Figure 1, private keys (A, B, C, D) are generated according to the *seed*, *index*, and *security level* [36]. The corresponding *addresses*, also called *leaves* (A', B', C', D'), can be generated respectively [37,38]. By applying the hash functions to narrow the addresses, the *root* of the Merkle tree can be obtained. In a MAM stream, a single MHT only lasts for a short period of time, each message contains the root of the next Merkle tree (or the future direction of the channel) [32].

Each message is signed with the one-time signature (OTS) scheme. Each leaf in the MHT corresponds to 1 OTS scheme. This means that each tree can produce the same number of messages as the number of leaves in the MHT [15].

In an MHT, the set of complementary hashes of a given leaf are the *siblings* of this leaf. As shown in Figure 1, the siblings of leaf *A'* (in red color) are *B"* and *Hash(C",D")*. By combining a given leaf and its siblings, the *root* of an MHT can be calculated.

A complete MAM transaction should include a signature section and the masked message section. The signature is created from one of the private keys corresponding to one of the leaves. The masked message consists of the raw data that need to be shared, the root of the next MHT, the index of the chosen leaf (branch index), and the siblings of this leaf. Figure 2 shows an example of a MAM stream with 2 transactions.

XSL•FO
RenderX

Figure 1. An example of Merkle Hash Tree with 4 leaves.

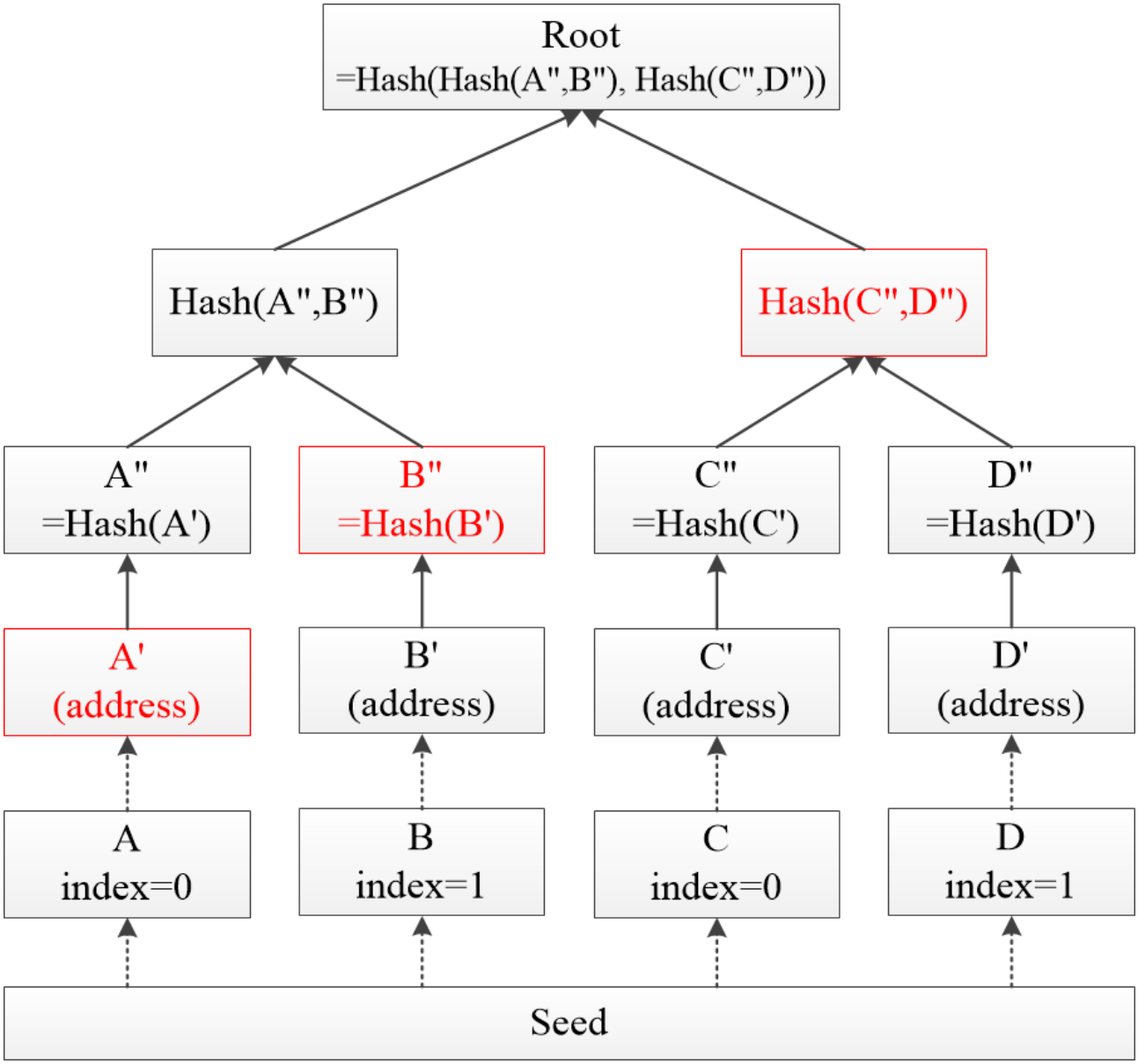
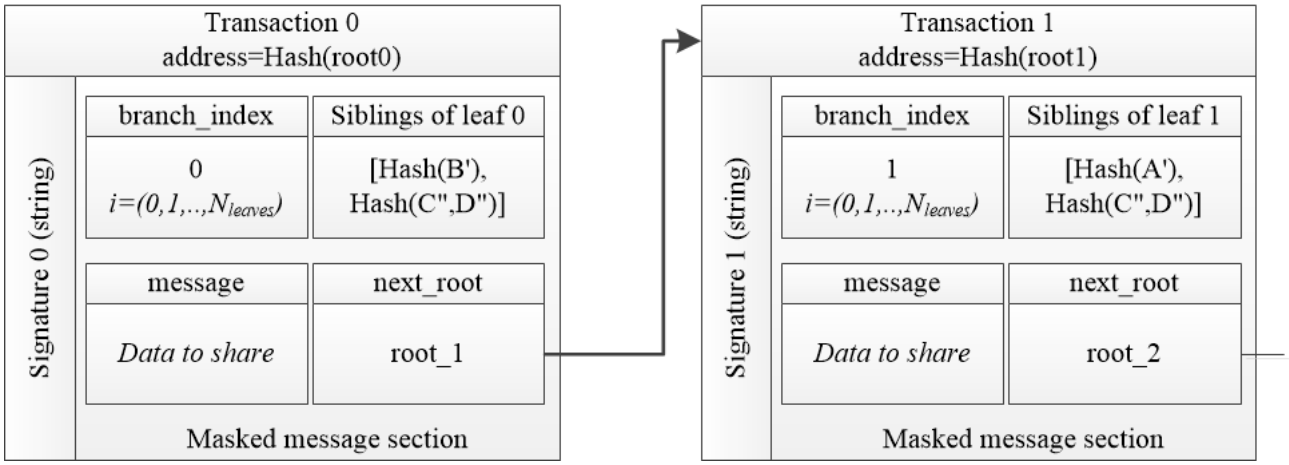


Figure 2. Structure of a Masked Authenticated Messaging stream with 2 transactions.



XSL•FO
RenderX

Privacy and Encryption Modes

MAM has 3 privacy and encryption modes to control the visibility and access of a channel: public, restricted, and private.

In public mode, the root of MHT is directly used as MAM transaction address and channel key. Therefore, any user who receives a message randomly or intentionally can then decode it by using the address of the message.

In private mode, the hash of the MHT root is used as the address, and the message is decrypted using the *root*. This prevents random users from decrypting the message if they stumble across it as they are unable to derive the *root* from the hash.

In restricted mode, an authorization key, named as *sideKey* in this study, is added based on private mode. The address used to attach to the network is the hash of the *sideKey* and the *root* (according to the current MAM source code [38], only the hash of root is used, which differs from the introduction of IOTA website [32,36]). It enables a message publisher to revoke access to future messages from subscribers by changing the *sideKey*.

To consume a MAM message, the receiver needs to use the *root* to calculate the address of the transaction and fetch the masked message. Then, use the *root*, and *sideKey* in restricted mode, to decrypt the masked message.

In a MAM channel, the current message contains the address of the following message, whereas the previous ones are not referenced. This adds the forward secrecy character to a channel. When users are authorized with the correct decryption key, they could follow a MAM stream from the current transaction, but there is no way to read previous messages.

Objective of This Paper

The objective of this study was to integrate IOTA Tangle with IoT to develop a health data sharing system, which could support secure, fee-less, tamper-resist, high-scalable, and granular-controllable health data exchange. The data source could include both wearable devices and stationary sensors in a smart environment such as smart home. The feasibility of the proposed system needs to be verified with a prototype system and its application in a practical case.

Methods

System Architecture

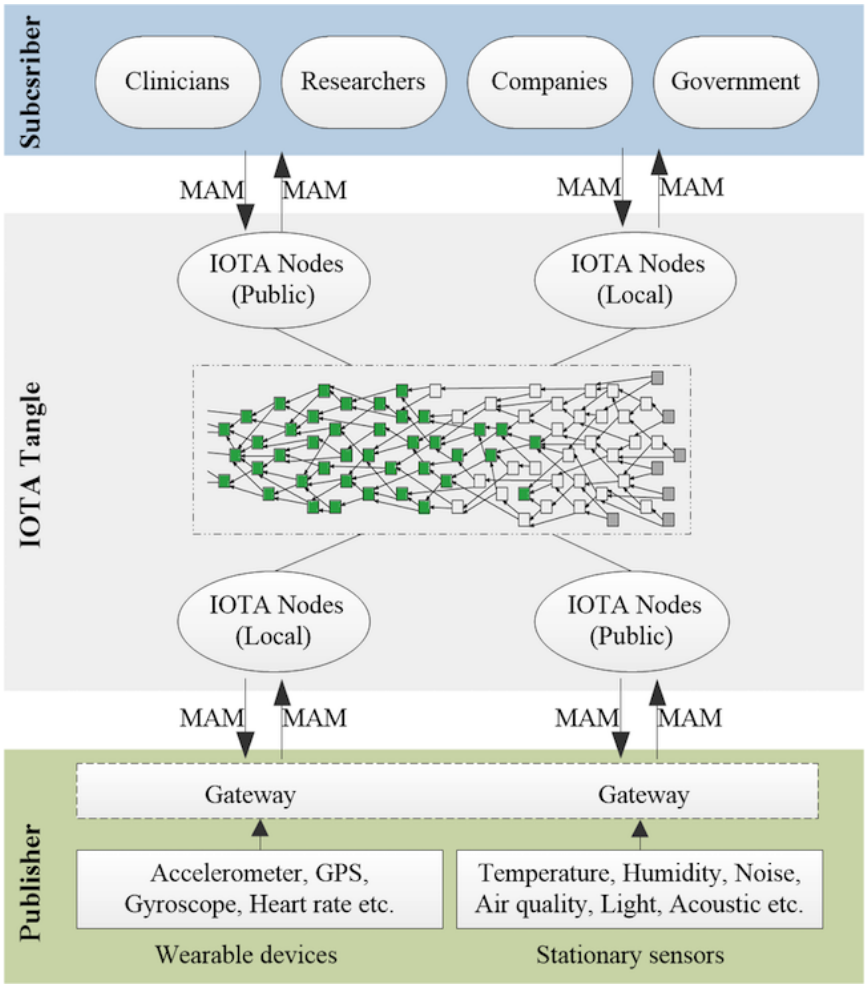
The architecture of the proposed health care data sharing system is presented in Figure 3. There are 2 roles involved in this system, data publisher and data subscriber. The publisher can be an individual, a family, or any other organization who possesses smart devices and sensors. These devices, sensors, and their owners produce health-related data, which are then published to the Tangle using specific encryption and privacy protocols. The data are published in their own channels, and each channel has an address. The subscribers of a data channel will receive the new published data. The published data are usually encrypted, and an extra decryption key may be necessary to decrypt the received data.

All the data are published and received through an IOTA node, which is a computer connected to the IOTA network. Users may use their own node or use public nodes. A user can be a data publisher and a subscriber at the same time. For example, a patient can publish his or her health data, and his or her doctor can subscribe these data and make evaluation accordingly. Afterwards, the doctor can publish the evaluation result to the Tangle, and the patient can subscribe this channel and receive the result.

Due to the limitations of size, power supply, and computing capability, most wearable devices and environmental sensors cannot publish or receive data directly to or from the Tangle. In this case, a gateway layer will be necessary, which could be a computer, a smartphone, or a single-board computer such as Raspberry Pi [39].

XSL•FO
RenderX

Figure 3. Architecture of the proposed health data sharing system based on IOTA Tangle. GPS: Global Positioning System; MAM: Masked Authenticated Messaging.



Implementation

To verify the feasibility of the proposed health data sharing system and demonstrate the implementation process, a prototype has been developed. The structure of the prototype is shown in Figure 4.

A portable human movement monitoring system using smartwatches was developed previously for the remote diagnosis of essential tremor (ET) [40,41]. The Pebble smartwatch [42] in this system could measure the triaxial acceleration data for tremor evaluation and activity recognition. The customized apps in the smartphone allows users to report their location, activity name, tremor level, self-evaluation about the disease, and other factors related to the disease, such as medication, alcohol, and coffee intake. These data, after integrated with other sensor data generated by the smartphone, will be compressed and uploaded to the remote server via internet for analysis using machine learning techniques.

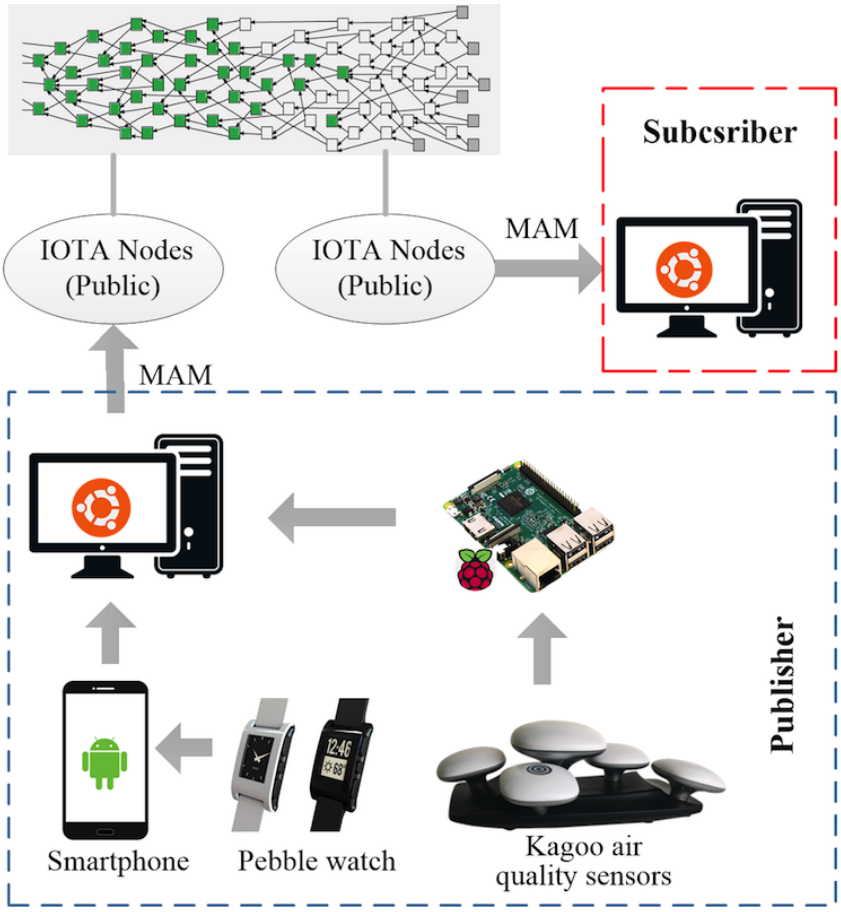
In addition to the movement monitoring system, we added an environmental monitoring system composed of Kagoo air quality sensors [43] and Raspberry Pi [39]. The Kagoo sensors could measure various environmental factors such as temperature, humidity, noise, and the content of pollutions in the air, including particulate matter, formaldehyde, total volatile organic components, benzene, carbon dioxide, carbon monoxide, ozone, and nitrogen dioxide. These sensors can be freely combined and plugged into a motherboard, which can communicate with the single-board computer Raspberry Pi through wired or Wi-Fi connection. A Python program running on Raspberry Pi could fetch and preprocess the environmental data from air sensors. More technical details, including the hardware manual and software codes, are openly accessible [44].

We use these 2 data collecting systems to represent wearable devices and stationary context sensors. The combination of these 2 data sources could provide a more complete understanding about users' health-related information.

In this prototype, the Pebble smartwatch, Android smartphone, and air quality sensors compose the sensing layer; the smartphone and Raspberry Pi play the role of gateway corresponding to the architecture of the proposed system. The data collecting frequency varies among different devices. The acceleration data from the smartwatch are recorded with a frequency of 25 Hz and uploaded every minute in a batch. The frequency of the data from smartphone depends on user's habit and usually is less than once per hour. The Kagoo sensors record environmental data once per minute.

XSL•FO
RenderX

Figure 4. Prototype of the health data sharing system using smartwatches, smartphone, air sensors, and Raspberry Pi. MAM: Masked Authenticated Messaging.



In this prototype, instead of directly published to the Tangle through the gateway as shown in Figure 1, the raw data are first sent to a local server for processing. The reasons are 2-fold. First, the frequency of the acceleration data from smartwatch is much higher than the other 2 data sources. Publishing these raw data to the Tangle will lead to a long lagging period. Therefore, on the server side, the acceleration data will pass through a tremor evaluation module based on deep learning approaches [41]. The output will be a shorter message per minute with a time stamp and a tremor score based on the classification result. This shorter message will be published to the Tangle. The raw acceleration data will be saved in a private database for future use. The second reason is to simplify the experiment of testing the average waiting time of publishing messages. The data from different sources are all gathered in the server and published through the same node in a concentrated period to obtain a more reliable result.

In terms of privacy and encryption modes, the environmental data are published in public MAM mode, whereas the patient report data and tremor evaluation data are published in restricted MAM mode.

To consume the published data over the Tangle, subscribers only need to know the address of the channel if the data are published in public mode, whereas an extra decryption key is needed for the data in restricted mode as introduced previously. Both data publishing and data receiving were realized through the JavaScript programs, which are introduced with details in the following experiment and results section.

Results

Experiment

An experiment was conducted to prove the feasibility of the proposed system, which can broadcast and receive combined health data from both wearable devices and stationary environmental sensors. In general, 3 types of data are tested, including tremor level based on smartwatch acceleration data, patient reports from smartphone, and environmental data from air sensors. The environmental data were broadcasted using public MAM protocol, whereas the other 2 types of data were broadcasted using restricted MAM protocol. The authentication keys of restricted mode were changed during a broadcast stream to demonstrate how a user could revoke access to the data they generate in future.

All the data were broadcasted and received in JSON format. For each type of data, 100 trials of broadcasting were realized to test the average waiting time. The data were published using a computer equipped with a 4-core Intel Core i5-4460 3.2 GHz CPU, a 12 GB of RAM memory, and the Ubuntu Linux 18 64-bit version operating system. The data were published through a public IOTA node [45,46]. The memory usage was 50%, and the number of neighbors was 12 when connected to this node during the experiment. The complete scripts for

XSL•FO
RenderX

publishing and receiving JSON data over the Tangle are openly available [47].

Experiment Outcomes

Figure 5 shows an example of published environmental data over the Tangle using public MAM mode. It displays that the *address* of the channel is the same as the *root* of the MHT. Any user who knows the *address* could fetch the message and decrypt it with *root*, which is the same to the *address*.

Figure 6 presents an example of patient report data published in restricted mode. In this case, the *address* is the hash of the MHT *root*, which is totally different. Subscribers need to know both the *address* and the extra encryption key (*side_key*) to fetch and decrypt the message. In restricted mode, the publisher can send a subscriber the *address* and *side_key* to grant him or her access to the current and future messages in the data stream. To revoke the authorization, the publisher just needs to change the *side_key* when publishing a new message, and subscribers without the new *side_key* will loss the access to this message and future ones.

The combination of public and restricted MAM protocols could provide users granular control over their heath data, which could bring great benefit to the health care system. For instance, in our prototype, when a patient wants to be diagnosed, he or she can share with the neurologist the *address* and the *side_key* to the report data and tremor evaluation data streams from a certain time. Afterwards, the neurologist will be able to fetch all 3 data streams as the environmental data are published in public mode. After the diagnosis, the patient can change the key to revoke the authorization, as shown in Figure 7.

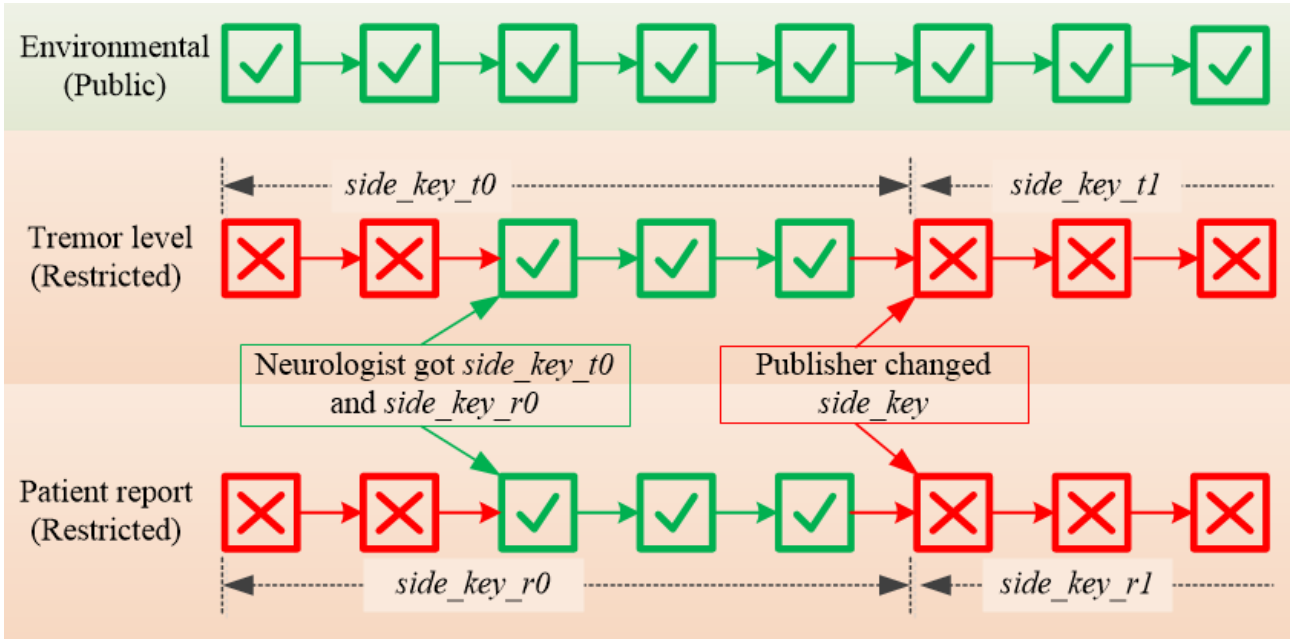
Figure 5. Environmental data published to the Tangle with public Masked Authenticated Messaging mode.

```
Root:  YCHCPSUSUSMSODZI9RESQDYGGBYGJCVLJHYQWAUGMECGZXQZDGQUDSFWAIBHWJBXLLAJDVSIKWFZOZDZE
Address:  YCHCPSUSUSMSODZI9RESQDYGGBYGJCVLJHYQWAUGMECGZXQZDGQUDSFWAIBHWJBXLLAJDVSIKWFZOZDZE
waiting_time:37852
location: Celsa Group Office, timestamp: 2019-01-09 00:02:00, pm2_5: 12.286, pm10: 13.143,t
voc: 0.036, co2: 0.2, temperature: 26.7, humidity: 14.913,illumination: 0.0, noise: 66.897,
 hcho: 0.02, co: 0, c6h6: 0.0, no2: 0, o3: 0
```

Figure 6. Patient report data published to the Tangle with restricted Masked Authenticated Messaging mode.

```
Root:  LCMNXLYG9MBPSFUZRLJPQO9YOQZAGWVWCQNNGWDGCXSTZAQGCJVGVWCOHYNPOGQICSWQIELBPLUEDVCUW
Address:  ZWSYJNEKDGWLOTBSJUKNLOVPOHFNUHBRILTOWDPKKKBXFRGDWKNTPAJGYIOHSCQQQMCVEVRECEXQWMPBB
waiting_time:33040
date: 2016-06-09T17:13:41.000Z, alcohol: No, caffee: No,Medicine: SumialMysoline, t1: 2016-
06-09T17:05:02.000Z, ArmExtendL: 4, ArmExtendR: 4,t2: 2016-06-09T17:06:03.000Z, TouchNoseL:
 4, TouchNoseR: 4, t3: 2016-06-09T17:06:52.000Z, Writting: 4, t4: 2016-06-09T17:08:39.000Z,
```

Figure 7. Granular access control over messages published over the Tangle by combing public and restricted Masked Authenticated Messaging protocol.



XSL•FO
RenderX

Table 1. Result of the data broadcasting experiment using Masked Authenticated Messaging (MAM) protocol.

Data	MAM mode	Size (bytes)	Waiting time per message (seconds)		
			Mean (SD)	Maximum	Minimum
Air quality	Public	260	20.41 (8.11)	55.41	8.81
Tremor level	Restricted	29	17.35 (5.37)	33.41	7.81
Personal report	Restricted	570	19.99 (7.61)	55.17	10.10
Air quality from Raspberry Pi	Public	260	22.56 (5.97)	40.36	12.87

The summary of the waiting time for publishing the 3 types of data to the Tangle based on 100 trials is presented in Table 1. The result shows that there is no obvious difference among these 3 types of data in terms of waiting time for publishing to the Tangle, although their message length and encryption modes are different. This is because of the fact that, in IOTA Tangle, the size of a transaction is 2673 trytes, which is about 1650 bytes. It means that as far as a message is shorter than this limit, the waiting time of publishing such messages should be similar on the same node. The actual waiting time depends on the computing capability of the node to perform proof of work, and it may vary from a few seconds to more than half minute according to our tests as shown in Table 1.

Currently, the bottleneck of the data publishing speed is the total number of nodes connected to the Tangle network and the condition of the specific node used by the publishing device. It is expected that the time to publish data from a local server or from the single board computer should be similar. Aiming to verify this consumption, an extra test using Raspberry Pi to publish the air quality data was conducted in addition to the experiment of publishing 3 types of data using a local server. The result is presented in Table 1. It shows that there is no obvious difference regarding the waiting time for publishing a message, which verified the aforementioned consumption.

Discussion

Principal Findings

This study explored the application of emerging distributed ledger technology in the health care domain. We proposed a health data sharing system by converging IoT, IOTA Tangle, and MAM protocol. It makes possible of a reliable marketplace for the individuals to share their health-related data with hospitals, researcher, industry companies, or any other organizations in a secure and controllable way. In return, individuals can get benefit from their own data in monetary, medical services, or other forms. On the other hand, researchers and companies will be able to gather relevant data for their studies, clinical trials, or product development.

Most existing studies about the applications of DLT and IOT in health care either focused on the conceptual design of health data sharing systems or discussed relevant policies from managerial perspectives. In comparison, this study not only proposed an application framework supported by DLT and IOT technologies but also implemented a prototype system in practice from technical perspective.

Through an experiment based on a prototype system, we demonstrated how the health-related data are collected and published to the Tangle in different encryption and privacy options. Our experiment showed that combining public and restricted MAM data streams, individuals are enabled to define granular access controls to different data consumers. The proposed system could facilitate the development of fee-less, secure, and efficient health data sharing marketplace to handle the big data generated by numerous IoT devices, and hence, pave way to the promising intelligent health care.

Although the current implementation of IOTA Tangle and its MAM protocol are already usable, they are still under development and are evolving rapidly. The current waiting time for attaching a message to the Tangle may vary from a few seconds up to more than 1 min. Although this is faster than other block-based protocols, there is still large room for performance improvements, as the more nodes connected to the Tangle network, the faster a transaction can be approved.

Limitations

The feasibility of the proposed health data sharing system using IOTA Tangle and MAM protocol was verified through the experiments based on a prototype system in a controlled environment. There are a few limitations worth to be mentioned.

First, in the prototype, a local server was introduced between the gateway layer and the IOTA nodes. The aim was to handle the large amount of raw acceleration data and simplify the testing process. In practical application, this local server can be excluded. The sensor data can be published to the Tangle directly from IoT devices or through a gateway such as smartphone or Raspberry Pi. This could enable the real machine-to-machine communication and make it easier for large scale implementation.

Another limitation of this pilot study is that a public IOTA node was used for publishing and receiving data over the Tangle. The lagging time varied depending on the workload of that public node, which is always stable. In practical implementation, a private node should be set up according to the practical requirements.

Conclusions

IOTA Tangle, together with the MAM communication protocol, could provide a fee-less, secure, highly scalable, and quantum-immune data sharing platform. The fast development of IoT is upgrading health care industry from digital to intelligent. The converging of IOTA Tangle, MAM, and IoT could significantly accelerate the health data sharing and pave way to realizing the vision of intelligent health care. The

XSL•FO
RenderX

proposed solution in this study overcomes many of the challenges faced by other traditional block-based solutions in terms of cost, efficiency, scalability, and flexibility in data access management. It could be applied in many scenarios of health care, such as remote diagnosis, chronic disease monitoring, and elderly caring, as introduced in the previous ET diagnosis experiment. Patients can publish their own health data to the Tangle with different encryption options and authorize medical experts to access to the tremor and activity data during a period. Experts can also share the diagnosis result with patients or their relatives.

This solution could be useful in many other areas such as rehabilitation, sports and fitness, and labor health protection in workplaces, which indicates the directions for future work. For example, wearable devices can be used to monitor workers' positions, activities, working load, and health indicators such as heart rate and blood pressure. Environmental sensors can be used to monitor the working conditions including the air quality, temperature, humidity, noise, and illumination. All these data or the periodical statistic results can be published to the Tangle and authorize access to different stakeholders such as Environmental Health and Safety experts, production managers, and government audit departments to better understand the health status of workers and avoid overfatigue or injuries.

Conflicts of Interest

None declared.

References

1. Gubbi J, Buyya R, Marusic S, Palaniswami M. Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gen Comput Syst 2013 Sep;29(7):1645-1660. [doi: 10.1016/j.future.2013.01.010]
2. Zheng X, Rodríguez-Monroy C. The development of intelligent healthcare in China. Telemed J E Health 2015 May;21(5):443-451. [doi: 10.1089/tmj.2014.0102] [Medline: 25671730]
3. Risteska Stojkoska B, Trivodaliev K. A review of Internet of Things for smart home: challenges and solutions. J Clean Product 2017 Jan;140:1454-1464. [doi: 10.1016/j.jclepro.2016.10.006]
4. Ni Q, García Hernando AB, Pau de la Cruz I. A context-aware system infrastructure for monitoring activities of daily living in smart home. J Sens 2016;2016:1-9. [doi: 10.1155/2016/9493047]
5. Son D, Lee J, Qiao S, Ghaffari R, Kim J, Lee JE, et al. Multifunctional wearable devices for diagnosis and therapy of movement disorders. Nat Nanotechnol 2014 May;9(5):397-404. [doi: 10.1038/nnano.2014.38] [Medline: 24681776]
6. Zheng X, Vieira A, Marcos S, Aladro Y, Ordieres-Meré J. Activity-aware essential tremor evaluation using deep learning method based on acceleration data. Parkinsonism Relat Disord 2019 Jan;58:17-22. [doi: 10.1016/j.parkreldis.2018.08.001] [Medline: 30122598]
7. Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. Industr Mngmnt Data Syst 2015 Oct 19;115(9):1704-1723. [doi: 10.1108/IMDS-03-2015-0087]
8. Varshney U. Pervasive healthcare and wireless health monitoring. Mobile Netw Appl 2007 Jul 12;12(2-3):113-127. [doi: 10.1007/s11036-007-0017-1]
9. Zhang J, Xue N, Huang X. A secure system for pervasive social network-based healthcare. IEEE Access 2016;4:9239-9250. [doi: 10.1109/ACCESS.2016.2645904]
10. Manyika JM. McKinsey & Company. By 2025, Internet of things applications could have $11 trillion impact URL: https://tinyurl.com/y4jcejlg [accessed 2019-02-01] [WebCite Cache ID 75rGpVMDl]
11. Peterson K, Deeduvanu R, Kanjamala P, Boles K. A blockchain-based approach to health information exchange networks. 2016 Presented at: NIST Workshop Blockchain Healthcare; 2016; Gaithersburg, MD URL: https://www.colleaga.org/sites/default/files/12-55-blockchain-based-approach-final.pdf
12. EUR-Lex: EU law. Brussels: European Parliament; 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46 URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 [accessed 2019-05-21] [WebCite Cache ID 78Y7eDQns]
13. Sønstebø D. IOTA Blog. IOTA Data Marketplace URL: https://blog.iota.org/iota-data-marketplace-cb6be463ac7f?gi=12ef095fee1f [accessed 2019-02-01] [WebCite Cache ID 75rH5vOj9]
14. Callegati F, Cerroni W, Ramilli M. Man-in-the-middle attack to the HTTPS protocol. IEEE Secur Privacy Mag 2009 Jan;7(1):78-81. [doi: 10.1109/MSP.2009.12]
15. Brogan J, Baskaran I, Ramachandran N. Authenticating health activity data using distributed ledger technologies. Comput Struct Biotechnol J 2018;16:257-266 [FREE Full text] [doi: 10.1016/j.csbj.2018.06.004] [Medline: 30101004]
16. Nakamoto S. ResearchGate. 2008. BitcoA peer-to-peer electronic cash system URL: https://www.researchgate.net/publication/228640975_Bitcoin_A_Peer-to-Peer_Electronic_Cash_System [accessed 2019-05-18] [WebCite Cache ID 78T2oB3dp]
17. Böhme R, Christin N, Edelman B, Moore T. Bitcoin: economics, technology, and governance. J Econ Perspect 2015 May;29(2):213-238. [doi: 10.1257/jep.29.2.213]

XSL•FO
RenderX

18. Ali ST, Clarke D, McCorry P. Bitcoin: Perils of an Unregulated Global P2P Currency. School of Computing Science Technical Report Series 2015 May 01:283-293 University of Newcastle upon Tyne.
19. Harlev M, Sun YH, Langenheldt K. Breaking Bad: De-Anonymising Entity Types on the Bitcoin BlockchainUsing Supervised Machine Learning. 2018 Presented at: 51st Hawaii International Conference on System Sciences; January 3-6, 2018; Hawaii p. 3497-3506 URL: https://core.ac.uk/download/pdf/143481278.pdf
20. Mamoshina P, Ojomoko L, Yanovich Y, Ostrovski A, Botezatu A, Prikhodko P, et al. Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. Oncotarget 2018 Jan 19;9(5):5665-5690 [FREE Full text] [doi: 10.18632/oncotarget.22345] [Medline: 29464026]
21. Ethereum F. Ethereum.org. 2018. URL: https://www.ethereum.org/dao [accessed 2019-02-01] [WebCite Cache ID 75rHBYTtW]
22. IOTA Blog. 2018. Meet the Tangle URL: https://www.iota.org/research/meet-the-tangle [accessed 2019-02-01] [WebCite Cache ID 75rHI9zXS]
23. Bitcoin. Confirmation URL: https://en.bitcoin.it/wiki/Confirmation [accessed 2019-02-01] [WebCite Cache ID 75rHLWc13]
24. Statista. 2017. Average confirmation time of Bitcoin transactions from January to December (in minutes) URL: https://www.statista.com/statistics/793539/bitcoin-transaction-confirmation-time/ [WebCite Cache ID 75rHcAn3d]
25. Blockchain. 2019. Transaction Rate URL: https://www.blockchain.com/en/charts/transactions-per-second [accessed 2019-02-01] [WebCite Cache ID 75rHlcZRI]
26. Etherscan. 2019. Ethereum transaction chart URL: https://etherscan.io/chart/tx [accessed 2019-02-01] [WebCite Cache ID 75rHqWE2T]
27. BitInfoCharts. 2019. Bitcoin Avg. Transaction Fee historical chart URL: https://bitinfocharts.com/comparison/bitcoin-transactionfees.html#1y [accessed 2019-02-01] [WebCite Cache ID 75rIcoxMm]
28. Popov S. Contentful. 2018 Apr 30. The Tangle URL: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf [accessed 2019-05-18] [WebCite Cache ID 78T3lkPA3]
29. BTC.com. 2019. Pool Distribution URL: https://btc.com/stats/pool?pool_mode=year [accessed 2019-02-01] [WebCite Cache ID 75rHwwivP]
30. Brassard G, Høyer P, Tapp A. Quantum cryptanalysis of hash and claw-free functions. Latin American Symposium on Theoretical Informatics: Springer; 1998 Presented at: LATIN'98: Theoretical Informatics; April 20–24, 1998; Campinas, Brazil URL: https://link.springer.com/chapter/10.1007/BFb0054319 [doi: 10.1007/BFb0054319]
31. IOTA Blog. 2018. What is IOTA? URL: https://www.iota.org/get-started/what-is-iota [accessed 2019-02-01] [WebCite Cache ID 75rI5aNS4]
32. Handy P. IOTA Blog. 2017. Introducing Masked Authenticated Messaging URL: https://blog.iota.org/introducing-masked-authenticated-messaging-e55c1822d50e?gi=97f611b532fb [accessed 2019-02-01] [WebCite Cache ID 75rIDpJCj]
33. Adam L. ImperialViolet. 2013. Hash based signatures URL: https://www.imperialviolet.org/2013/07/18/hashsig.html [WebCite Cache ID 75rII7DYN]
34. Merkle R. A certified digital signature. Conference on the Theory and Application of Cryptology: Springer; 1989 Presented at: CRYPTO'89; August 20-24, 1989; Santa Barbara URL: https://link.springer.com/chapter/10.1007/0-387-34805-0_21
35. Merkle R. A digital signature based on a conventional encryption function. Conference on the theory and application of cryptographic techniques: Springer; 1987 Presented at: Conference on the Theory and Application of Cryptographic Techniques; August 16-20, 1987; Santa Barbara p. 369-378 URL: https://link.springer.com/chapter/10.1007/3-540-48184-2_32
36. ABmushi. Medium. IOTA: MAM Eloquently Explained URL: https://medium.com/coinmonks/iota-mam-eloquently-explained-d7505863b413 [accessed 2019-02-01] [WebCite Cache ID 75rIKrSNp]
37. ABmushi. Medium. IOTA: Signature and Validation URL: https://medium.com/@abmushi/iota-signature-and-validation-b95b3f9ec534 [accessed 2019-02-01] [WebCite Cache ID 75rIOTiI8]
38. GitHub. IOTA URL: https://github.com/iotaledger [accessed 2019-02-01] [WebCite Cache ID 75rIRmm28]
39. Raspberry Pi. URL: https://www.raspberrypi.org/ [accessed 2019-02-01] [WebCite Cache ID 75rI8dVyO]
40. Zheng X, Vieira Campos A, Ordieres-Meré J, Balseiro J, Labrador Marcos S, Aladro Y. Continuous monitoring of essential tremor using a portable system based on smartwatch. Front Neurol 2017;8:96 [FREE Full text] [doi: 10.3389/fneur.2017.00096] [Medline: 28360883]
41. Zheng X, Vieira A, Marcos S, Aladro Y, Ordieres-Meré J. Activity-aware essential tremor evaluation using deep learning method based on acceleration data. Parkinsonism Relat Disord 2019 Jan;58:17-22. [doi: 10.1016/j.parkreldis.2018.08.001] [Medline: 30122598]
42. FitBit. 2016. The Original Uncomplicated Pebble Classic URL: https://www.pebble.com/pebble-smartwatch-features [accessed 2019-04-15] [WebCite Cache ID 77eeWUHKU]
43. Circulate Health Science And Technology. 2019. URL: http://www.circulateoao.com/ [accessed 2019-02-01] [WebCite Cache ID 75rIWNn2b]
44. Sun S. GitLab. 2019. Indoor air quality project URL: https://gitlab.com/sunshengjing/iaq [WebCite Cache ID 75rLl4HKH]
45. IOTA. 2019. Public nodes list URL: https://www.tangle-nodes.com/?sorts[load]=1&sorts[tls]=-1 [WebCite Cache ID 75rNVsSJV]

XSL•FO
RenderX

46. IOTA Nodes List. URL: https://pow.iota.community:443 [accessed 2019-05-22] [WebCite Cache ID 78YaiXKKJ]
47. Zheng X. GitHub.: Github repository; 2019. iota mam data sharing URL: https://github.com/zhengxiaochen/iota_mam_data_sharing [WebCite Cache ID 75rNHxYIF]

Abbreviations

DLT: distributed ledger technologies
ET: essential tremor
IoT: Internet of Things
MAM: Masked Authenticated Messaging
MHT: Merkle Hash Tree
OTS: one-time signature

Edited by K Clauson, P Zhang; submitted 01.02.19; peer-reviewed by C Reis, E Minvielle, A Bello-Garcia; comments to author 31.03.19; revised version received 16.04.19; accepted 27.04.19; published 06.06.19.

Please cite as:
Zheng X, Sun S, Mukkamala RR, Vatrapu R, Ordieres-Meré J
Accelerating Health Data Sharing: A Solution Based on the Internet of Things and Distributed Ledger Technologies
J Med Internet Res 2019;21(6):e13583
URL: https://www.jmir.org/2019/6/e13583/
doi:10.2196/13583
PMID:31172963

©Xiaochen Zheng, Shengjing Sun, Raghava Rao Mukkamala, Ravi Vatrapu, Joaquín Ordieres-Meré. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Designing a Distributed Ledger Technology System for Interoperable and General Data Protection Regulation–Compliant Health Data Exchange: A Use Case in Blood Glucose Data

David Hawig[1*], MSc; Chao Zhou[1*], MD, DPhil; Sebastian Fuhrhop[1]; Andre S Fialho[1], MSc, DPhil; Navin Ramachandran[2], MBBS, MRCP, FRCR

[1]Pact Care BV, Amsterdam, Netherlands

[2]Centre for Health Informatics & Multiprofessional Education, University College London, London, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Navin Ramachandran, MBBS, MRCP, FRCR
Centre for Health Informatics & Multiprofessional Education
University College London
Radiology Department, University College Hospital
235 Euston Road
London, NW1 2BU
United Kingdom
Phone: 44 20 3447 9070
Fax: 44 20 3447 9297
Email: navinramachandran@nhs.net

Abstract

Background: Distributed ledger technology (DLT) holds great potential to improve health information exchange. However, the immutable and transparent character of this technology may conflict with data privacy regulations and data processing best practices.

Objective: The aim of this paper is to develop a proof-of-concept system for immutable, interoperable, and General Data Protection Regulation (GDPR)–compliant exchange of blood glucose data.

Methods: Given that there is no ideal design for a DLT-based patient-provider data exchange solution, we proposed two different variations for our proof-of-concept system. One design was based purely on the public IOTA distributed ledger (a directed acyclic graph-based DLT) and the second used the same public IOTA ledger in combination with a private InterPlanetary File System (IPFS) cluster. Both designs were assessed according to (1) data reversal risk, (2) data linkability risks, (3) processing time, (4) file size compatibility, and (5) overall system complexity.

Results: The public IOTA design slightly increased the risk of personal data linkability, had an overall low processing time (requiring mean 6.1, SD 1.9 seconds to upload one blood glucose data sample into the DLT), and was relatively simple to implement. The combination of the public IOTA with a private IPFS cluster minimized both reversal and linkability risks, allowed for the exchange of large files (3 months of blood glucose data were uploaded into the DLT in mean 38.1, SD 13.4 seconds), but involved a relatively higher setup complexity.

Conclusions: For the specific use case of blood glucose explored in this study, both designs presented a suitable performance in enabling the interoperable exchange of data between patients and providers. Additionally, both systems were designed considering the latest guidelines on personal data processing, thereby maximizing the alignment with recent GDPR requirements. For future works, these results suggest that the conflict between DLT and data privacy regulations can be addressed if careful considerations are made regarding the use case and the design of the data exchange system.

(J Med Internet Res 2019;21(6):e13665) doi:10.2196/13665

KEYWORDS

distributed ledger technology; directed acyclic graph; IOTA; IPFS; blockchain; Masked Authenticated Messaging, MAM; mobile health; blood glucose; diabetes; FHIR

XSL•FO
RenderX

Introduction

The delivery of high-quality health care requires the efficient and effective exchange of patient data [1,2]. To that end, health care systems across the globe have invested heavily over the past decade in a variety of solutions and tools that aim to improve the interoperable exchange of data, including health information exchanges, direct messaging, community clouds, and open application programming interfaces (APIs) [3]. These solutions have achieved different degrees of success in the real world, largely due to common concerns regarding privacy and security [4].

Distributed ledger technology (DLT) is a new type of database technology in which databases are connected in a distributed fashion through a peer-to-peer network and maintained by a consensus protocol [5]. Proponents of DLT believe that it offers great potential to make the exchange of health information immutable and secure [6-8]. However, the immutable and transparent character of these technologies may conflict with data privacy regulations or data processing best practices.

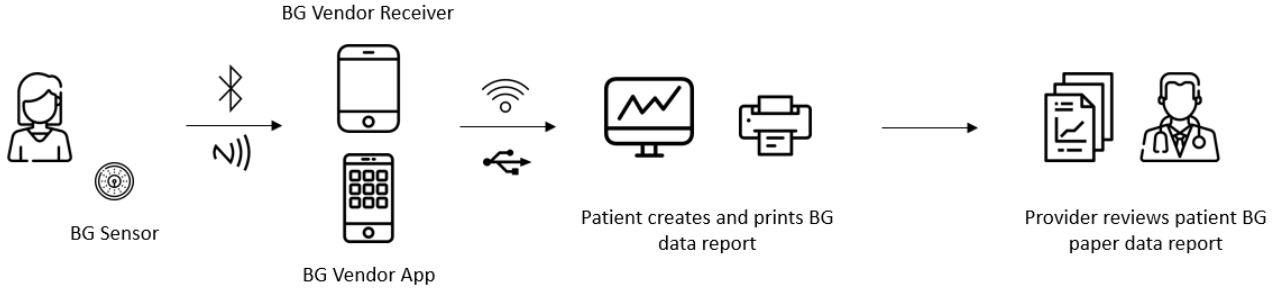
The European Union has recently instituted the General Data Protection Regulation (GDPR), which regulates the collection, processing, and securing of personal data, including protected health information (PHI). This new regulation sets out the way by which personal data are to be protected as well as defining the main rights of the person to whom the data relates (ie, the data subject): right of rectification, right to erasure, right of access, and rights related to automated processing. As detailed by the European Union Blockchain Observatory and Forum [9], in principle, there are no contradictions between the goals of GDPR and DLT. However, there are three areas in which GDPR still does not offer enough clarity about how real-world DLT applications should be developed. These areas include (1) accountability and roles (eg, how to identify a data controller in a public DLT), (2) anonymization of personal data (eg, what techniques are sufficient to anonymize personal data to the point where the resulting output can potentially be stored in a DLT), and (3) GDPR rights conflicts (eg, how to rectify or remove personal data that are recorded in a DLT that is immutable by nature, or who is responsible for requesting and managing the "freely, specific, informed, and unambiguous" consent from a data subject, especially if the data controller is not specified) [10].

With regards to anonymization of personal data, it is clear that GDPR does not apply to anonymized data and that this type of information can be stored on the DLT. However, what qualifies as anonymized is still not clear. The only indication today is that it must be irreversibly impossible to identify an individual through any of the means "reasonably likely to be used" [11]. Within the health care context, achieving this irreversible identification is even more difficult as PHI includes not only general personal data but also health status information, genetic data, and biometric data. In this way, GDPR sets a new bar for health care data anonymization compared to the "pseudonymization" methods currently used in clinical research, and where confidentiality is ensured through a simpler key coding of the data.

At this stage, the guidance provided by GDPR on how to process personal data simply refers to the need to minimize both the risk of reversal (eg, risk of reversing and reconstituting the original data) and the risk of linkability (eg, risk of linking anonymized data to an individual by examining patterns of usage or context, or by comparison to other pieces of information) [12,13]. A variety of techniques may accomplish this data anonymization, including obfuscation, encryption, hashing, and aggregation [14]. However, it remains unclear from a legal perspective, which of them (individually or in combination) are most adequate to convert personal data into anonymous data [9].

In this way, given the importance of minimizing the risk of data reversal and linkability for compliance with existing data privacy regulations, this study explores the potential of one specific DLT in supporting health information exchange by developing a proof-of-concept system for immutable, automated, and secure exchange of patient health information. Specifically, we focused on the patient-provider exchange of blood glucose (BG) data, not only because of the large number of patients affected by diabetes worldwide but also due to the importance of patient-provider communication of this data for improved treatment management. At present, patients usually track their glucose levels manually or rely on the vendor's software to compile a report (Figure 1). This manual process is time-consuming and error-prone [15], highlighting the need for systems supporting data exchange between self-monitoring BG devices and electronic health records (EHRs) in a secure, effective, and tamper-proof way.

Figure 1. Typical steps for current patient-provider exchange of continuous blood glucose (BG) data: (1) an adhesive patch holding the BG sensor is attached to the patient's skin and measures glucose readings in interstitial fluid throughout the day and night; (2) the sensor sends real-time readings wirelessly to a receiver/smart device app, so the user can view the information; (3) the receiver or smart device app displays current and historical glucose levels and allows for this data to be printed and/or exported (eg, .txt file); and (4) the patient and provider review together the paper notes or exported files.



XSL•FO
RenderX

Methods

System Design

We aimed to develop a proof-of-concept system for patient-provider exchange of glucose data (Figure 2) according to the following specifications: (1) patient-controlled, (2) fully digital, (3) interoperable, and (4) distributed logging and storage of data. *Patient-controlled* means that patients, in addition to providing consent for exchange, can also grant data access to selected parties. *Digital* specifies that data are exchanged in a digital format (from the BG remote sensor acquisition to the upload into a physician's EHR), minimizing any potential manual data entry errors. *System interoperability* specifies that the data generated by the BG remote sensor can be read, if consent is provided by the patient, by another information system from different health care parties (leveraging the Fast Healthcare Interoperability Resources [FHIR] standard). *Distributed logging and storage of data* means that all the data generated by the BG sensor are stored in a tamper-proof way across multiple distributed nodes instead of a single, centralized data repository.

Our system is built on four main modules: (1) a data conversion module which converts raw BG data from the device sensor into FHIR standard records, (2) a data processing module that transforms the generated FHIR records into a format that minimizes the risk of reversal and linkability, (3) a data storage and logging module which uploads the transformed FHIR records on a DLT, and (4) a key exchange module that allows establishing a patient-provider communication channel for the FHIR records to be exchanged.

The actual integration of the data from the DLT into a physician's EHR is out of the scope of this work because it is dependent on the EHR system and integration engine. However, the use of the FHIR standard in the proposed designs should make this a relatively straightforward process. Additionally, we assume the use case in which physicians will pull the data from the system on a periodic quarterly basis and not on a continuous, real-time basis. This sampling choice is based on the current patient-provider workflows and guidelines across several health care systems (eg, in Germany, diabetic patients have quarterly visits with physicians to review and update the disease management plan).

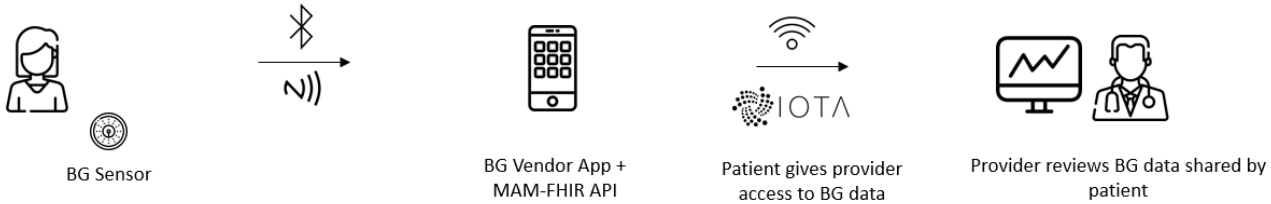
Also out of the scope of this work is the issue of data cache when patients are offline or in cases of network connections issues. In general, there are two points where data can be cached: (1) at the BG sensor and/or (2) on the mobile device used to read the data from the sensor. However, the exact architecture of the data cache implementation is highly dependent on the sensor and device manufacturer. Usually, these manufacturers leverage different libraries that provide access to the different secure storage solutions (eg, for Samsung Galaxy devices it would be the ARM's TrustZone).

Given the ongoing debates around what techniques are suitable for PHI anonymization, and because there currently is not a standard design for patient-provider data exchange, we propose two different variations for our proof-of-concept system. Both variations aim at being compliant with GDPR and attempt to minimize the risk of reversal and linkability in their own way, with their own advantages and disadvantages. These two variations consist of the same four modules but differ in the sense that one is solely based on a public distributed ledger named IOTA, which we call "public IOTA," and the second combines the IOTA public distributed ledger with an additional private distributed file system called InterPlanetary File System (IPFS), which we call "public IOTA plus private IPFS." A summary of these two variations can be seen in Figure 3, and a more detailed description of each module is provided subsequently.

Data Conversion Module

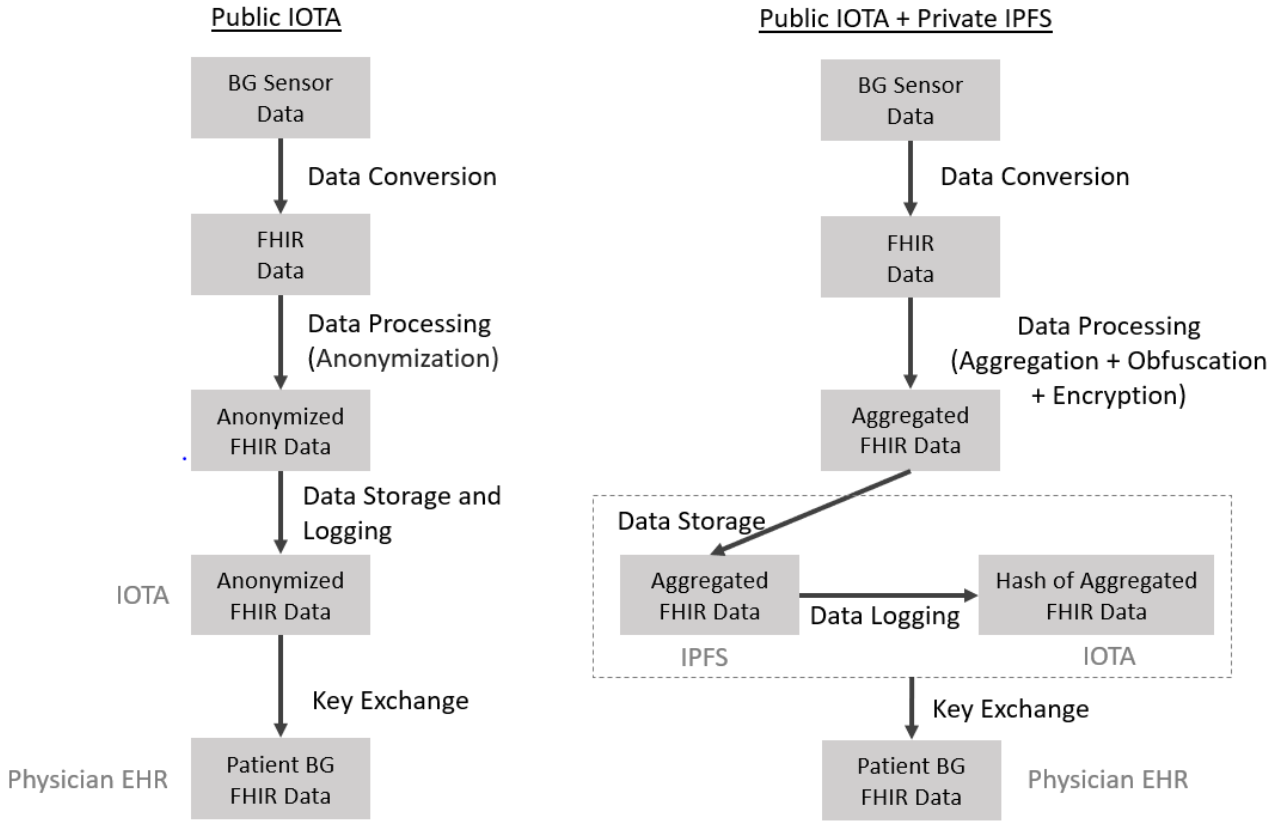
This module, common to both designs, converts raw data from the continuous BG system to the FHIR standard. FHIR was created by HL7 with the purpose of facilitating the interoperable exchange of health care-related data between different health care systems to make it easy to provide health care information to providers and individuals on a wide variety of devices and to allow third-party developers to provide medical apps that can be easily integrated into existing systems [16-18].

Figure 2. Proposed steps for a distributed ledger technology (DLT)-based patient-provider exchange of blood glucose (BG) data: (1) an adhesive patch holding the BG sensor is attached to the patient's skin and measures glucose readings in interstitial fluid throughout the day and night; (2) the sensor sends real-time readings wirelessly to a smart device app, so the user can view the information; (3) the smart device app displays current and historical glucose levels and is connected to an application programming interface (API; "MAM-FHIR API") that allows for these data to be exported to a DLT; and (4) if a patient provides consent, the interoperable data stored on the DLT can be automatically exported to a physician's electronic health record so that they can be reviewed.



XSL•FO
RenderX

Figure 3. Proof-of-concept system for patient-provider exchange of blood glucose (BG) data with two variations: (1) public IOTA and (2) public IOTA plus private IPFS (InterPlanetary File System). EHR: electronic health record; FHIR: Fast Healthcare Interoperability Resources.



The FHIR-based records are built from a set of modular components called "Resources" that have standard, agreed-on data elements with consistent meaning across sharing entities. All resources share a common set of features including (1) resource identity and metadata, (2) a human-readable XHTML summary, (3) a URL that identifies the resource, and (4) a set of defined data elements (a different set for each type of resource).

For ease of use and standardization purposes, we used a GATT-compliant Bluetooth continuous glucose monitoring system (Dexcom G4 PLATINUM) [19]. This device generates one BG measurement every 10 minutes. Converting the data from this continuous glucose monitoring system to an FHIR record was done by filling an FHIR observation blueprint for each generated glucose data measurement (mmol/L). The output of this module is one FHIR JavaScript Object Notation (JSON) file per each BG measurement.

Data Processing Module

Public IOTA

An FHIR record contains fields where PHI, such as patient name (eg, "John Doe"), date of birth (eg, "24-Sept 1932″), or medical record number (eg, "123456″) may be present. For purposes of GDPR adherence in this system design, because FHIR records will be stored on the public IOTA DLT, data anonymization was performed. Through this module, we have simply removed any existent PHI entries from each FHIR JSON file. For naming purposes, we will refer to this processed data as *anonymized FHIR records*.

Public IOTA Plus Private IPFS

With this design, all FHIR records were stored on a private distributed file system (IPFS), and the resulting hash was logged into the public IOTA DLT. For purposes of GDPR compliance and to minimize the risk of data reversal and linkability, we applied a sequential combination of steps in this module: (1) data aggregation that combines multiple FHIR JSON records into one single JSON file, (2) data obfuscation based on the JavaScript Obfuscator using the proposed medium settings that are applicable for a JSON file [20], and (3) data encryption of the obfuscated JSON file using the AES256-GCM function with a tag length of 128, which is also recommended by the German Medical Association and adopted by the US government [21].

The sampling used in the aggregation step was based on the periodic quarterly visits assumed previously for the patient-provider interaction. This aggregation window is, however, a variable that can be easily set for different time intervals (eg, real time, daily, weekly). For naming purposes, we will refer to this processed data as *aggregated FHIR records.*

Data Storage and Logging Module

Public IOTA

As shown in Figure 3, this design variation was based on a public DLT protocol only. We selected IOTA's DLT because it offers zero transaction fees and handles large transaction throughputs well. These properties make this technology particularly well-suited for data exchange across health care devices and systems [22]. What differentiates IOTA from blockchain-based DLTs is that transactions, instead of being

XSL•FO
RenderX

grouped into blocks and stored in sequential chains, are linked together in a so-called Tangle (Figure 4). This design allows for a validation process that takes form as a Web structure referred to as a directed acyclic graph rather than a linked list as is the case in blockchain-based DLTs. Transactions on the Tangle, therefore, can get issued simultaneously, synchronously, and continuously. To issue a transaction on the Tangle, one must validate two other transactions. Therefore, the more participants use the system, the more transactions are validated, making it highly scalable. This "pay-it-forward" validation process also obviates the need for financial rewards to incentivize participants.

To share, store, and retrieve encrypted data, we used IOTA's Masked Authenticated Messaging (MAM) module. This module encrypts messages (masking), confirms source origin (authentication), and creates a continuous message stream on the Tangle until the source stops publishing it (messaging). In a MAM stream, each message holds (1) the data, (2) a reference to the address of the next message only flowing in one direction (forward), and (3) a signature that proves that the publisher created that message. Given that unique IDs are created for each channel (known as "roots"), only those parties who are authorized are able to read and reconstruct the entire message stream.

To approve a transaction that is sent to the Tangle via MAM, computational resources based on Proof of Work (PoW) algorithms are used to find the answer to a simple cryptographic puzzle. These algorithms and the underlying peer-to-peer protocols use low processing resources, making them well-designed for small devices (eg, sensors) [23]. PoW can either be done locally on the device itself or externally via an IOTA node or a special API.

To test the performance of sending IOTA transactions via MAM, we assessed the time required to both create and attach 300 anonymized FHIR records. The choice for 300 messages was based on the intent to compare the performance of our system with previous research [22].

The message was created on the local hardware, whereas attachment of the message was performed via an external API. To create the message locally, we used a dummy C# Xamarin mobile phone app [24] that uses the Tangle.Net.Mam Nuget to generate MAM transactions [25] together with our FHIR code [26], running on a Samsung Galaxy S8 phone with 2.3 GHz Quad-Core Exynos M2 Mongoose. For the attachment, we used the external PoW via an API to Powsrv, which uses the "PiDiver" hardware [27] and uses a random selection of healthy IOTA nodes on the IOTA Mainnet to take different server response times into account. The use of this external API speeds up the sending process and reduces the device power consumption [27].

Public IOTA Plus Private IPFS

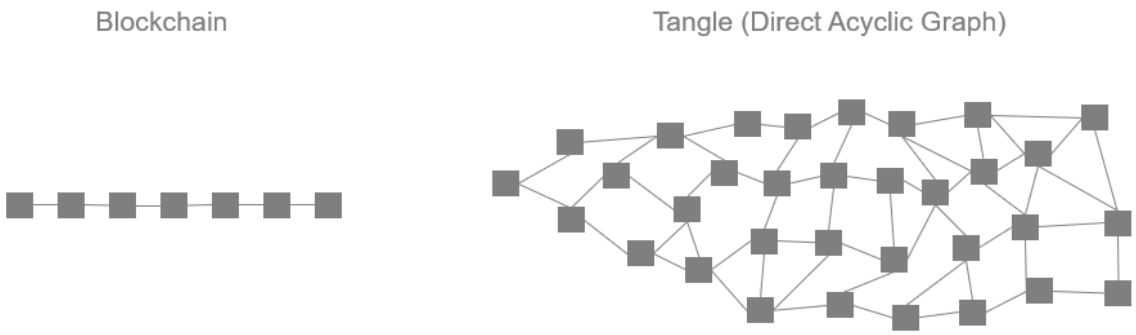
For the second design of our system, we included a distributed file storage system (IPFS) to account for two key regulatory considerations: (1) uncertainty with regards to what anonymization techniques are legally sufficient to transform PHI into anonymized data and (2) to minimize both the reversal and linkability risks.

The IPFS is a peer-to-peer distributed file system based on content-addressed hyperlinks. As such, it takes files and manages them based on their content, storing them and tracking their version using a generalized Merkle directed acyclic graph. These Merkle trees, or hash trees, allow secure verification of the contents of large data structures, using cryptographic hash functions that map data of arbitrary sizes to data of a fixed size (hash).

Advantages of this technology include (1) data stored on IPFS are not automatically distributed between all participants and only shared in the case of a request, (2) IPFS nodes are able to delete specific data at any given point in case of a request, and (3) it is easy to prove whether an input will result in a given hash, but incredibly difficult to derive the input from a hash [28].

In this design, the aggregated JSON files containing multiple individual FHIR records were uploaded into a private IPFS cluster via a writable IPFS gateway. Every participant of this network was publicly known and can be held accountable in the case of noncompliance with a data deletion request from the data subject. Therefore, this private setup allowed for the specification of the number of backup copies in the network, and for the definition of automatic rules, such as when to delete data in the case of a patient request. A previous study provides more details on how to set up these specific rules on an IPFS cluster [29].

Figure 4. Sequential block-based transactions of a blockchain (left) and IOTA directed acyclic graph-based transactions (right).



XSL•FO
RenderX

To link IPFS transactions to the authenticated and undeletable MAM transactions, hashes of the IPFS content were then shared via a MAM stream using the previously described IOTA libraries. For simplicity, we used the SHA256-256 hash function with Base58 encoding, which is the default hash function of IPFS.

The performance of this setup was assessed in two parts: (1) the time required to upload and generate a hash of the aggregated FHIR records using a writable IPFS gateway and (2) the time required to send this hash to the IOTA Tangle using MAM. This second test was also repeated 300 times using the same setup as in the previous design.

Key Exchange Module

The key exchange module allows the establishment of a patient-provider communication channel for the FHIR records to be exchanged. In this key exchange process, two parties exchange cryptographic keys, allowing them to exchange encrypted messages exclusively. The design of this module is dependent on the data logging and storage layer; therefore, two variations were also used in this work.

Public IOTA

In this design, an initial exchange of private data in person or via a different non-DLT method was required (to maintain GDPR compliance). We assumed an initial in-person key exchange where, during the first visit to the physician's office, the patient shared his or her channel keys using their mobile phone's near field communication (NFC) chip (tapped on a receiver that the physician owns, linked to the EHR). The information included in this key exchange includes the MAM root and channel key (Figure 5). When channel keys are exchanged with a physician, the physician can then retrieve and authenticate the anonymized BG monitoring device data stream(s) that reside on the Tangle, in a similar JSON format to Figure 6. You may notice that the ID of the JSON is pseudonymized with the first 64 letters of the MAM root (Figure 5), which is used to help the physician to uniquely identify the patient.

Figure 5. The Masked Authenticated Messaging (MAM) root and channel key of the in-person key exchange.

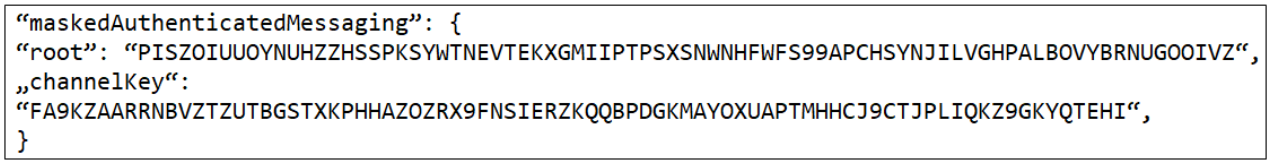
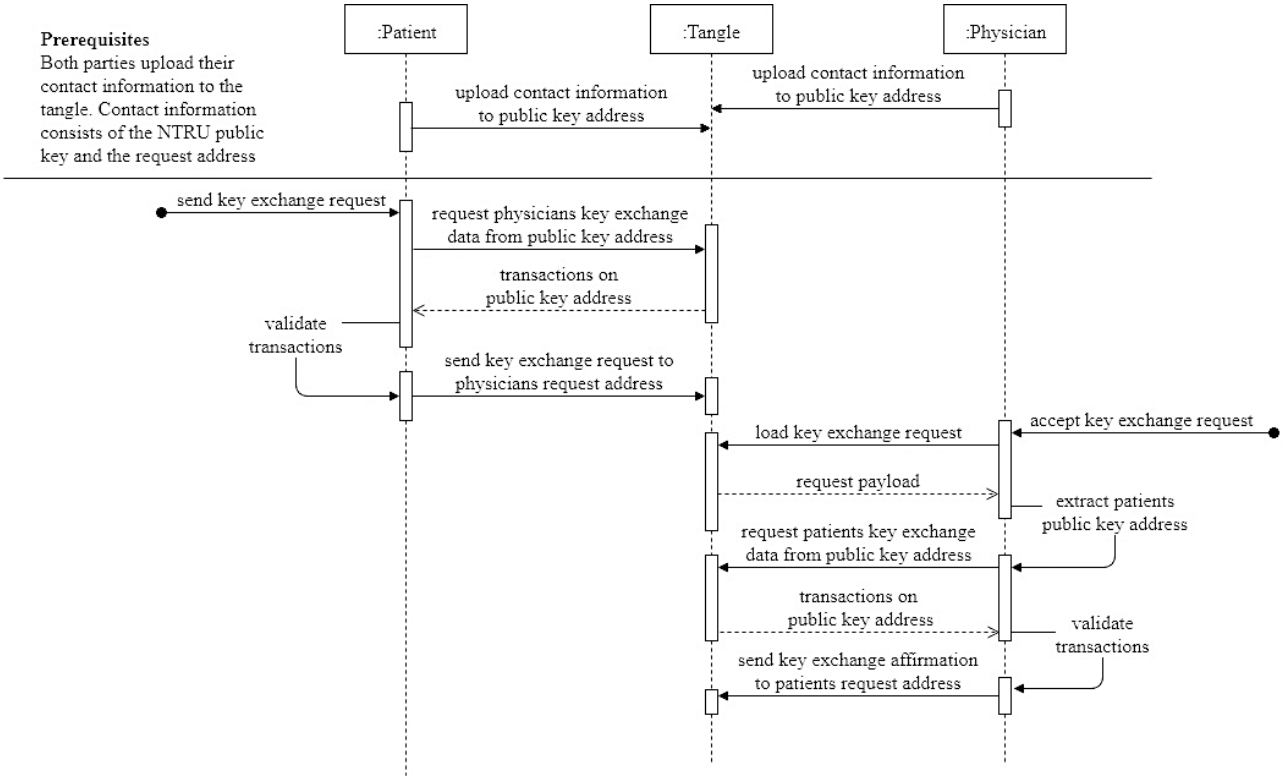


Figure 6. Anonymized JSON FHIR (JavaScript Object Notation Fast Healthcare Interoperability Resources) record stored on the public IOTA ledger with a pseudonymized ID consisting of the first 64 letters of the Masked Authenticated Messaging (MAM) root.

XSL•FO
RenderX

Figure 7. Steps for a remote key exchange via a public key exchange. NTRU: Nth degree - truncatedpolynomial ring units.



At any point in time, patients can revoke access to their data stream by simply updating their MAM channel's authorization key. If multiple doctors have access to the same channel, the parties who should continue to have access would get the new channel key through an additional key exchange process.

Public IOTA Plus Private IPFS

In this design, the initial exchange of keys and personal data can either take place in person during the first visit to the physician's office or remotely via a public key exchange. Although the exchange of keys in person is the most secure option, there are use cases in which it is more convenient or not possible to exchange information unless it is done remotely (eg, follow-up remote consultation or providing data access to researchers). For this variation, we used a remote public key exchange system which assumes that both patient and physician have their contact information (an Nth degree - truncated polynomial ring units [NTRU] public key uploaded to their public key IOTA address along with a link to their request address) published to the Tangle.

As shown in Figure 7, the steps for key exchange in this case are (1) the patient sends a key exchange request to the physician's request address, (2) the physician decrypts this request with his NTRU private key (due to the nature of NTRU, only he can decrypt the data), (3) the physician confirms the key exchange request by sending his details to the patient's request address, and (4) the channel access data can now be generated from the portions sent in steps 1 and 3, and both parties can now send encrypted data to the secured channel's address.

Results

We assessed (1) the performance of the data processing module, (2) the time required to send and store FHIR records in a DLT system by comparing the performance of the data storage and logging module for each of the two system designs, and (3) the time required for the remote key exchange process.

Data Processing

For the public IOTA design, after the anonymization step, the size of the JSON file containing one BG FHIR record decreased from 2509 bytes to 857 bytes. During the aggregation step of the public IOTA plus private IPFS design, we combined the BG FHIR records for the equivalent of one-quarter of patient monitoring. This resulted in a JSON file with an approximate size of 34 MB. After this aggregation, and after also applying the JavaScript Obfuscator, the file size increased to 64 MB. Finally, after applying the AES256-GCM encryption, the resultant file had a size of 93 MB. It is important to note that the encryption time depended both on the encryption algorithm and processing hardware used. In this study, using the AES256-GCM algorithm and a Samsung Galaxy S8 device (2.3Ghz Quad-Core Exynos M2 Mongoose), it took less than 0.1 seconds to encrypt a 93 MB file which, in the larger picture of total transaction time, is negligible.

Data Storage and Logging

Public IOTA

Table 1 shows the time required to both create and attach an anonymized FHIR record (with 857 bytes) with MAM. The total time required to create and attach a single message was mean 6.1 (SD 1.9) seconds (create: mean 3.5, SD 1.1 seconds; attach: mean 2.5, SD 0.8 seconds).

XSL•FO
RenderX

Table 1. Transaction times for storing and logging anonymized records of Fast Healthcare Interoperability Resources on the public IOTA design and public IOTA plus private InterPlanetary File System (IPFS) design and remote key exchange times.

Design and action	Trials	Time (ms), mean (SD)	Time (ms), range	Variance (ms^2)
Public IOTA				
Create	300	3525 (1182)	2042-8100	1,397,997
Attach	300	2545 (765)	1357-8923	584,728
Public IOTA + private IPFS				
Create	300	3636 (1371)	2249-12,673	18,794,477
Attach	300	3522 (576)	2161-5146	331,554
Remote key exchange				
Send request	10	5160 (1801)	3100-8500	3,247,111
Accept request	10	5790 (1253)	4000-8500	1,572,111

Public IOTA Plus Private IPFS

Next, we looked at the time required to upload and generate a hash of the aggregated FHIR record (93 MB) using a writable IPFS gateway. The hashing of this file was almost instantaneous; therefore, the total time required was fully dependent on the user's bandwidth speed to upload this file. Our upload speed was on average 3 MB/s, which led to a total upload time of approximately 31.0 (SD 8.5) seconds. After uploading the FHIR records on the IPFS gateway, a hash was returned with a 46-character format similar to QmP543pymsKVHUdMg YQzSRbG7HoDSrVajhVRfrbtvhnGAQ. The mean times required to send this hash to the IOTA Tangle using MAM are presented in Table 1.

Using this public IOTA plus private IPFS design, the total amount of time required to share an aggregated quarter of glucose data was mean 38.1 (SD 13.4) seconds (upload: mean 31.0, SD 8.5 seconds; create: mean 3.6, SD 1.4 seconds; attach: mean 3.5, SD 0.6 seconds).

Key Exchange

With the remote setup described in the public IOTA plus private IPFS design, we were able to exchange keys remotely and establish a secure connection in mean 10.9 (SD 3.1) seconds (send request: mean 5.2, SD 1.8 seconds; accept request: mean 5.8, SD 1.3 seconds) via the IOTA Tangle (Table 1).

Discussion

Principal Findings

Systems based on DLT hold great potential to help improve health information exchange. However, their design needs to comply with the latest regulations (eg, GDPR) to minimize the risk of inappropriate processing and/or storage. In this study, we assessed the potential of one specific DLT (IOTA) for the exchange of BG data, by designing and developing a proof-of-concept system considering important regulatory considerations. Such systems should also meet satisfactory performance and usability from the user perspective; therefore, we tested the performance of our design using a variety of measures.

Given the current open questions around data anonymization, an optimal system does not exist at this stage. Therefore, we designed two variations. Each of these have their own pros and cons, including (1) data reversal risk, (2) data linkability risks, (3) processing time, (4) file size compatibility, and (5) overall system complexity. This assessment is summarized in Table 2.

Reversal Risk

Reversal risk represents the risk of being able to reconstitute the original data from anonymized or modified data. In the public IOTA design, given that FHIR records uploaded into the Tangle were stripped of any personal data, this risk is not applicable. For the public IOTA plus private IPFS design, the hash that is uploaded into the Tangle is the result of multiple processing layers (aggregation, obfuscation, and encryption). Reverse engineering this hash into the original file is, computationally, a massive task since it requires trialing the immense number of possible combinations of inputs, which can range from a few bytes to hundreds of terabytes in size. Even in the future, with the natural advancements in computing, it is difficult to imagine that it will be possible to extract the data from one 46-letter hash, which stands for 93 MB of obfuscated and encrypted data.

XSL•FO
RenderX

Table 2. Summary of the advantages and disadvantages of the two variations of the proposed proof-of-concept.

Feature	Public IOTA	Public IOTA + private IPFS[a]
Reversal risk	N/A[b]	Low
Linkability risk	Medium	Low
Processing time	Low	Low
File size compatibility	Small files	Any file size
Complexity	Low	Medium

[a]IPFS: InterPlanetary File System.

[b]N/A: not applicable.

Linkability Risk

The linkability risk represents the risk that it is possible to link anonymized data to an individual by examining patterns of usage or context, or by comparison to other pieces of information. In the public IOTA design, by tracking the frequency of the uploads of data into the Tangle, plus reading the actual BG levels, it may be possible to link this information to a particular individual. A possible solution could entail the encryption of the raw BG levels. However, in the light of GDPR, any encryption algorithm is susceptible of being reversed in the future [9]; therefore, encrypted raw data can potentially be decrypted and the actual glucose levels shown. Given the cases of reidentification of anonymized health records reported in the literature [13], one should be careful not to assume that anonymized PHI can be uploaded to a public DLT without any risk. The exchange of multiple keys during the first appointment between patient and provider could be another solution to reduce the amount of PHI per MAM stream and therefore reduce the linkability risk in this design. This solution has not been explored in this work.

For the public IOTA plus private IPFS design, each resulting hash is unique, so there is no obvious way to cross-analyze the data and therefore determine who these data belong to.

Data Processing Times

The message transaction times obtained were in line with previous research [22]. The overall low transaction times for messages on the IOTA DLT (on average between 6.1 to 38.1 seconds) suggest that both designs can be used for a continuous data stream. It is important to note here that, by definition, one MAM message actually consists of three transactions, which in theory could be performed in parallel. With the implementation of this code optimization (dividing the average times shown in Table 1 by 3), the actual average time per message creation and attachment would be 1175 ms and 848 ms, respectively. This would result in a total of approximately 2 seconds per IOTA message sent via MAM.

With hardware acceleration [27] of PoW, it is possible to achieve a PoW time of 300 ms. The overall response time of the external API was 848 ms, suggesting a further API bottleneck of approximately 548 ms. To achieve faster transaction times, users could set up their own hardware-accelerated IOTA nodes rather than using the centralized API service. Increasing the number of users running their own PoW also increases the overall health of the network, as it increases its decentralized character.

Specifically for the public IOTA plus private IPFS design, because the hashing itself is almost instant, the main limiting factor is the upload speed of the network. For this design, as a general recommendation, we recommend uploading of large files to the IPFS gateway only while connected to a fast WLAN network.

Finally, it is also relevant to note that for providers to read the BG data, taking into account the use case of quarterly consultations, the public IOTA design will benefit from using a “prefetching” of this quarterly data into the provider’s EHR. For the public IOTA plus private IPFS design, given the relatively low reading times (available on request), we do not envision the need for prefetching.

File Size Compatibility

In the public IOTA design, the time required to create and send a message via MAM fully depends on the size of the file. Larger files may therefore not be suitable to be exchanged using this design. In the public IOTA plus private IPFS design, because the hashes are usually much smaller than the initial data and they can be used to identify the data itself, we can conclude that this design is well-suited for the exchange of larger files.

Complexity

Regarding the overall design complexity, the public IOTA design is significantly simpler compared with the public IOTA plus private IPFS design. Two features, in particular, make the second design more complex: (1) the multiple steps involved in the data processing module require additional processing that can lead to larger file sizes (eg, using the quantum-secure NTRU encryption resulted in a file size of 821 MB) and (2) the setup of the private IPFS cluster requires additional implementation work.

The use case chosen in this paper focused on the remote exchange of BG measurements between a patient and a physician, where the physician “reads” the data on a periodic basis during the consultation, to review and update the diabetes management plan. Nevertheless, we are also confident that either design could be used in the use case of continuous streaming of individual BG measurements (ie, real-time data exchange and monitoring), as well in other types of remote health data exchange (such as the use cases proposed by Cohen et al [30]).

XSL•FO
RenderX

Related work in this field falls into two categories: (1) use of blockchain to improve health care data access and interoperability, and (2) assessment of how blockchain can be used considering data privacy regulations.

One of the first works describing the use of blockchain to tackle interoperability barriers was authored by the MedRec team [31]. In this work, a permissioned blockchain network based on a proof-of-work incentive was proposed to facilitate EHR data sharing and authentication. Following this work, Zhang et al [32] took a step further and described how EHR data could be securely and scalably shared to improve collaborative clinical decision support. Specifically, an FHIR-based smart contract system was proposed for exchanging health data, in which the blockchain stores encrypted metadata and an off-chain solution is used for clinical data.

This work does present an FHIR-based provider-provider solution, although our work proposes a complementary FHIR-based solution for patient-provider data exchange. On the topic of FHIR-based interoperability, Peterson et al [33] described a blockchain solution using a new type of consensus mechanism: an FHIR-based "Proof of Interoperability." This approach is unique in the sense that it takes security and interoperability as the central tenet of its core design.

In contrast to the previous studies, it is important to note two additional articles positioning blockchain as a tool for patient-provider data exchange instead of provider-provider exchange. Ichikawa et al [34] developed a mHealth system using a mobile phone app that enables a patient-provider exchange of information but for the particular goal of insomnia cognitive behavioral therapy. Balsari et al [35] proposed a use case leveraging the high mobile phone penetration and availability of unique ID systems in India to facilitate health data exchange between more than 500 million Indian citizens and their providers.

The use of blockchain in light of data privacy regulations has been previously described by Zyskind et al [36] across three domains: data ownership, data transparency and auditability, and access control. This article describes a method similar to our second proposed design, namely an on-chain solution combined with off-blockchain storage to construct a personal data management platform focused on privacy. Conversely, Al Omar et al [37], arguing that decentralized approaches for data exchange may fail to ensure overall privacy, applied additional cryptographic functions and data processing procedures to encrypt patients' data and to ensure pseudonymity.

Limitations

In this work, we address some of the points of tension between GDPR and DLT by proposing two designs that enable the exchange of PHI using this technology. However, future work needs to address in more detail the remaining points of conflict, such as how to identify a data controller in a public DLT ecosystem formed by multiple health care stakeholders or how to collect and manage an individual's express consent for the processing of health data.

It is also important to note a few limitations with the designs proposed in this study. First, using a patient's mobile phone's NFC chip for key exchange is not currently practical because it requires every physician's office to have an NFC reader connected to the facility's EHR. Alternatives to this could be an in-person key exchange method or secure local wireless network. Second, in our testing, the continuous glucose monitoring device generated a data point every 10 minutes, meaning that significant battery usage could be expected over the day to create all FHIR records. The actual implications on the device battery may constitute an important limitation requiring further investigation.

Finally, it is also important to point out that the experiments run in this study were carried out on a Samsung Galaxy S8 only. The CPU of this device was used to calculate the resource hashes locally; therefore, it can be expected that different mobile devices will require different times to create and attach an FHIR resource.

Conclusion

The design of a DLT-based system for health data exchange needs to take into careful consideration the respective use case. In this paper, we proposed and developed two possible designs that aim to be compliant with recent data privacy regulations, minimizing any risks of misappropriate data processing and returning satisfactory performance and usability. One design was based solely on the public distributed ledger IOTA and the second used IOTA plus a private IPFS cluster. Our findings suggest that the first design is simpler to implement but requires special attention to minimize the risk of personal data linkability, and that the second design allows for the exchange of larger files at the expense of higher complexity.

Acknowledgments

We thank the IOTA Foundation for the technical support provided throughout the elaboration of this work.

Conflicts of Interest

ASF, CZ, DH, and SF received support from the IOTA Foundation, a German nonprofit organization, which aims to support initiatives that explore the use of the IOTA DLT into solving real-world problems. No financial compensation was provided for the actual designing and drafting of this paper. NR is a member of the IOTA Foundation and holds a portfolio of cryptocurrencies, including IOTA. No payment has been made for designing and drafting this paper.

References

XSL•FO
RenderX

1. Health Information Technology Policy Committee. Report to Congress: Challenges and Barriers to Interoperability. 2015 Dec. URL: http://www.healthit.gov/facas/sites/faca/files/HITPC_Final_ITF_Report_2015-12-16%20v3.pdf [accessed 2019-01-28] [WebCite Cache ID 75l9W84Ts]
2. Office of the National Coordinator for Health Information Technology (ONC). Report to Congress: Report on Health Information Blocking. 2015 Apr. URL: https://www.healthit.gov/sites/default/files/reports/info_blocking_040915.pdf [accessed 2019-01-28] [WebCite Cache ID 75l9yzMzF]
3. Dean AA. Journal of AHIMA. 2018 Sep 01. The future of healthcare data exchange URL: https://journal.ahima.org/2018/09/01/the-future-of-healthcare-data-exchange/ [accessed 2019-01-28] [WebCite Cache ID 75lAFBlps]
4. Mello MM, Adler-Milstein J, Ding KL, Savage L. Legal barriers to the growth of health information exchange-boulders or pebbles? Milbank Q 2018 Mar;96(1):110-143. [doi: 10.1111/1468-0009.12313] [Medline: 29504197]
5. Gordon WJ, Wright A, Landman A. NEMJ Catalyst. 2017 Feb 09. Blockchain in health care: decoding the hype URL: https://catalyst.nejm.org/decoding-blockchain-technology-health/ [accessed 2019-01-28] [WebCite Cache ID 75lBGztfz]
6. Gordon WJ, Catalini C. Blockchain technology for healthcare: facilitating the transition to patient-driven interoperability. Comput Struct Biotechnol J 2018 Jun 30;16:224-230 [FREE Full text] [doi: 10.1016/j.csbj.2018.06.003] [Medline: 30069284]
7. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 01;24(6):1211-1220. [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
8. Angraal S, Krumholz HM, Schulz WL. Blockchain technology: applications in health care. Circ Cardiovasc Qual Outcomes 2017 Sep;10(9):1-3. [doi: 10.1161/CIRCOUTCOMES.117.003800] [Medline: 28912202]
9. European Union Blockchain Observatory and Forum. Blockchain and the GDPR. 2019. URL: https://www.eublockchainforum.eu/sites/default/files/reports/20181016_report_gdpr.pdf?width=1024&height=800&iframe=true [accessed 2019-01-28] [WebCite Cache ID 75lBR9uoP]
10. Official Journal of the European Union. 2016 Apr 14. Regulation (EU) 2016/679 of the European Parliament and of the council URL: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679 [accessed 2019-01-28] [WebCite Cache ID 75lBks9Uo]
11. Working Party on the Protection of Individuals with Regard to the Processing of Personal Data. Opinion 05/2014 on Anonymisation Techniques. 2014 Apr 10. URL: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf [accessed 2019-01-28] [WebCite Cache ID 75lByGg8V]
12. Ohm P. Broken promises of privacy: responding to the surprising failure of anonymization. UCLA Law Rev 2010;57(6):1701-1777 [FREE Full text]
13. El Emam K, Jonker E, Arbuckle L, Malin B. Correction: a systematic review of re-identification attacks on health data. PLoS One 2015;10(4):e0126772. [doi: 10.1371/journal.pone.0126772] [Medline: 25880057]
14. McCallister E, Grance T, Scarfone K. Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). Gaithersburg, MD: National Institute of Standards and Technology; 2010 Apr 06. URL: https://ws680.nist.gov/publication/get_pdf.cfm?pub_id=904990 [accessed 2019-01-28] [WebCite Cache ID 75lCnhAD7]
15. Golden SH, Sapir T. Methods for insulin delivery and glucose monitoring in diabetes: summary of a comparative effectiveness review. J Manag Care Pharm 2012 Aug;18(6 Suppl):S1-S17. [Medline: 22984955]
16. Health Level Seven International (HL7). Introducing HL7 FHIR URL: https://www.hl7.org/fhir/summary.html [accessed 2019-01-28] [WebCite Cache ID 75lD9Ggq4]
17. Kawamoto K, Hongsermeier T, Wright A, Lewis J, Bell DS, Middleton B. Key principles for a national clinical decision support knowledge sharing framework: synthesis of insights from leading subject matter experts. J Am Med Inform Assoc 2013 Jan 01;20(1):199-207 [FREE Full text] [doi: 10.1136/amiajnl-2012-000887] [Medline: 22865671]
18. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Sep;23(5):899-908 [FREE Full text] [doi: 10.1093/jamia/ocv189] [Medline: 26911829]
19. Bluetooth SIG. 2014. Continuous glucose monitoring URL: https://www.bluetooth.com/wp-content/uploads/Sitecore-Media-Library/Gatt/Xml/Services/org.bluetooth.service.continuous_glucose_monitoring.xml [accessed 2019-01-28] [WebCite Cache ID 75lDK7qtt]
20. Kachalov T. Github. 2018. JavaScript obfuscator source code URL: https://github.com/javascript-obfuscator/javascript-obfuscator [accessed 2019-05-31] [WebCite Cache ID 78mTr2bvb]
21. Nechvatal J, Barker E, Bassham L, Burr W, Dworkin M, Foti J, et al. Report on the Development of the Advanced Encryption Standard (AES). J Res Natl Inst Stand Technol 2001 Jun;106(3):511-577. [Medline: 27500035]
22. Brogan J, Baskaran I, Ramachandran N. Authenticating health activity data using distributed ledger technologies. Comput Struct Biotechnol J 2018 Jul 17;16:257-266. [doi: 10.1016/j.csbj.2018.06.004] [Medline: 30101004]
23. Popov S. The Tangle. 2018 Apr 30. URL: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf [accessed 2019-01-28] [WebCite Cache ID 75lE21ycq]
24. Pact Care BV. Github. 2018. Chiota Source code URL: https://github.com/PACTCare/Chiota [accessed 2019-05-26] [WebCite Cache ID 78f8j4UP4]
25. Felandil S. Nuget. Tangle.Net.Mam URL: https://www.nuget.org/packages/Tangle.Net.Mam/ [accessed 2019-05-26] [WebCite Cache ID 78f99boVw]

XSL•FO
RenderX

26. Pact Care BV. Github. 2018. PACT FHIR source code URL: https://github.com/PACTCare/Pact.Fhir [accessed 2019-05-26] [WebCite Cache ID 78f9DrxzT]
27. Pototschnig T. Microengineer. 2018 Apr 25. IOTA PoW hardware accelerator FPGA for Raspberry Pi (und USB) URL: https://microengineer.eu/2018/04/25/iota-pearl-diver-fpga/ [accessed 2019-01-28] [WebCite Cache ID 75lEJoicf]
28. Benet J. IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3). 2019. URL: https://ipfs.io/ipfs/QmR7GSQM93Cx5eAg6a6yRzNde1FQv7uL6X1o4k7zrJa3LX/ipfs.draft3.pdf [accessed 2019-01-28] [WebCite Cache ID 75lERb3GH]
29. Protocol Labs. IPFS. 2019. IPFS cluster overview URL: https://cluster.ipfs.io/documentation/overview/ [accessed 2019-01-28] [WebCite Cache ID 75lEUsqWO]
30. Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from Project HealthDesign. JMIR Hum Factors 2016 Oct 19;3(2):e26 [FREE Full text] [doi: 10.2196/humanfactors.5919] [Medline: 27760726]
31. Azaria A, Ekblaw A, Vieira T, Lippman A. MedRec: using blockchain for medical data access and permission management. In: Proceedings of the 2nd International Conference on Open and Big Data. 2016 Aug Presented at: 2016 2nd International Conference on Open and Big Data; Aug 22-24, 2016; Vienna, Austria p. 25-30.
32. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST. FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 2018 Jul 29;16:267-278. [doi: 10.1016/j.csbj.2018.07.004] [Medline: 30108685]
33. Peterson K, Deeduvanu R, Kanjamala P, Boles K. ONC/NIST. 2016. A Blockchain-Based Approach to Health Information Exchange Networks URL: https://www.healthit.gov/sites/default/files/12-55-blockchain-based-approach-final.pdf [accessed 2019-01-29] [WebCite Cache ID 75n9xR1cf]
34. Ichikawa D, Kashiyama M, Ueno T. Tamper-resistant mobile health using blockchain technology. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e111 [FREE Full text] [doi: 10.2196/mhealth.7938] [Medline: 28747296]
35. Balsari S, Fortenko A, Blaya JA, Gropper A, Jayaram M, Matthan R, et al. Reimagining health data exchange: an application programming interface-enabled roadmap for India. J Med Internet Res 2018 Jul 13;20(7):e10725 [FREE Full text] [doi: 10.2196/10725] [Medline: 30006325]
36. Zyskind G, Nathan O, Pentland A. Decentralizing privacy: using blockchain to protect personal data. In: IEEE Secur Priv Work. 2015 Presented at: IEEE Symposium on Security and Privacy Workshops; May 18-20, 2015; San Jose, CA p. 180-184.
37. Al Omar A, Bhuiyan MZ, Basu A, Kiyomoto S, Rahman MS. Privacy-friendly platform for healthcare data in cloud based on blockchain environment. Future Gener Comp Sy 2019 Jun 08;95:511-521. [doi: 10.1016/j.future.2018.12.044]

Abbreviations

API: application programming interface
BG: blood glucose
DLT: distributed ledger technology
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
GDPR: General Data Protection Regulation
IPFS: InterPlanetary File System
JSON: JavaScript Object Notation
MAM: Masked Authenticated Messaging
NFC: near field communication
NTRU: Nth degree - truncated polynomial ring units
PHI: protected health information
PoW: Proof of Work

Edited by P Zhang, K Clauson; submitted 08.02.19; peer-reviewed by W Gordon, S Balsari; comments to author 27.04.19; revised version received 09.05.19; accepted 13.05.19; published 14.06.19.

Please cite as:
Hawig D, Zhou C, Fuhrhop S, Fialho AS, Ramachandran N
Designing a Distributed Ledger Technology System for Interoperable and General Data Protection Regulation–Compliant Health Data Exchange: A Use Case in Blood Glucose Data
J Med Internet Res 2019;21(6):e13665
URL: http://www.jmir.org/2019/6/e13665/
doi:10.2196/13665
PMID:31199293

XSL•FO
RenderX

©David Hawig, Chao Zhou, Sebastian Fuhrhop, Andre S Fialho, Navin Ramachandran. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Potential of Blockchain Technology for Health Information Exchange: Experimental Study From Patients' Perspectives

Pouyan Esmaeilzadeh[1*], PhD; Tala Mirzaei[1*], PhD

Department of Information Systems and Business Analytics, College of Business, Florida International University, Modesto A Maidique Campus, Miami, FL, United States

[*]all authors contributed equally

Corresponding Author:
Pouyan Esmaeilzadeh, PhD
Department of Information Systems and Business Analytics
College of Business
Florida International University, Modesto A Maidique Campus
11200 SW 8th Street
Miami, FL, 33199
United States
Phone: 1 (305) 348 3302
Email: pesmaeil@fiu.edu

Abstract

Background: Nowadays, a number of mechanisms and tools are being used by health care organizations and physicians to electronically exchange the personal health information of patients. The main objectives of different methods of health information exchange (HIE) are to reduce health care costs, minimize medical errors, and improve the coordination of interorganizational information exchange across health care entities. The main challenges associated with the common HIE systems are privacy concerns, security risks, low visibility of system transparency, and lack of patient control. Blockchain technology is likely to disrupt the current information exchange models utilized in the health care industry.

Objective: Little is known about patients' perceptions and attitudes toward the implementation of blockchain-enabled HIE networks, and it is still not clear if patients (as one of the main HIE stakeholders) are likely to opt in to the applications of this technology in HIE initiatives. Thus, this study aimed at exploring the core value of blockchain technology in the health care industry from health care consumers' views.

Methods: To recognize the potential applications of blockchain technology in health care practices, we designed 16 information exchange scenarios for controlled Web-based experiments. Overall, 2013 respondents participated in 16 Web-based experiments. Each experiment described an information exchange condition characterized by 4 exchange mechanisms (ie, direct, lookup, patient-centered, and blockchain), 2 types of health information (ie, sensitive vs nonsensitive), and 2 types of privacy policy (weak vs strong).

Results: The findings show that there are significant differences in patients' perceptions of various exchange mechanisms with regard to patient privacy concern, trust in competency and integrity, opt-in intention, and willingness to share information. Interestingly, participants hold a favorable attitude toward the implementation of blockchain-based exchange mechanisms for privacy protection, coordination, and information exchange purposes. This study proposed the potentials and limitations of a blockchain-based attempt in the HIE context.

Conclusions: The results of this research should be of interest to both academics and practitioners. The findings propose potential limitations of a blockchain-based HIE that should be addressed by health care organizations to exchange personal health information in a secure and private manner. This study can contribute to the research in the blockchain area and enrich the literature on the use of blockchain in HIE efforts. Practitioners can also identify how to leverage the benefit of blockchain to promote HIE initiatives nationwide.

(J Med Internet Res 2019;21(6):e14184) doi:10.2196/14184

KEYWORDS

health information exchange; patients; privacy; trust; risk; perception

XSL•FO
RenderX

Introduction

Health Information Exchange Models

Individuals usually seek health care services from several providers who may practice in either affiliated or unaffiliated institutions. Accordingly, without a systematic connection among providers, patients' medical information can become fragmented, outdated, and incomplete in health care organizations [1]. Health information exchange (HIE) is a data exchange mechanism that was introduced and prompted by the Health Information Technology for Economic and Clinical Health Act in 2009 to improve care coordination among health care providers and reduce medical errors [2]. HIE refers to the process of electronic transfer of patient health information and medical data among health care providers and institutions [3]. Interoperability associated with HIE initiatives requires electronic communication among organizations to ensure that patient medical records in one health care organization are seamlessly incorporated into another.

Different sharing mechanisms are being used by public and private health care organizations to facilitate information exchange initiatives [4]. Existing studies in HIE indicate that the following 3 exchange models are mainly applied by health care entities to electronically transmit patient health information: (1) direct, (2) query-based, and (3) patient-centered exchange [5]. In the direct model, a provider can share encrypted patient medical records with a known recipient [6]. This exchange model facilitates point-to-point data exchange in which the sender is aware of the recipient's identity and patients' medical records can be exchanged directly from one health care organization to another via widely adopted email protocols. Direct exchange initiatives, which are principally based on trust between providers, incorporate medical records into the recipient's electronic health record (EHR) system or clinical inbox in a secure network governed by health care entities. The direct model is able to improve communication and coordination among health care organizations involved in providing treatments by securely exchanging identifiable information of patients.

The query-based models (lookup systems) grant health care providers the ability to find and request information on a patient from other providers. In this exchange mechanism, a central repository plays a critical role where electronic medical records are aggregated from multiple health care organizations' EHR systems and will be stored in a hub [7]. Thus, the requesting health care organizations are able to use a lookup process to pull required information from the data storage pool [8]. The query-based model is mainly designed to create a mechanism to efficiently provide relevant, aggregated, and cross-organizational health records for care quality measurement and disease registries development.

The last model refers to a patient-centered exchange mechanism in which medical records related to episodes of care are transmitted from providers to patients. For instance, patients are able to view the laboratory results, radiology reports, progress notes, and medications that are uploaded on patient portals after each visit and share such records with other health care entities as required [9]. This exchange architecture is developed to enable patients to engage in their care process, manage their health information, and become a component of data-sharing efforts by considering a mediating role for them. Patients can leverage the patient-centered HIE models, which are designed and controlled by health care institutions, to reinforce their access and control over their own health records.

Role of Patients in Health Information Exchange

Given the huge amount of information exchanged among health care organizations, patients would rely on HIEs to improve treatment process, enhance care coordination, and increase the quality of care before they actually experience the possible effects [10]. In this setting, risk can also arise because patients may be concerned that too much personal information is shared or erroneous health information is exchanged among health care providers through HIEs [11]. In the HIE context, patients may not directly share their health information through exchange mechanisms, and they are distant from care providers who actually use these systems. However, patients are recognized as an important beneficiary of HIE projects because their consent is required for sharing their health information [12]. Patients are also considered as a significant producer of health information and their attitudes toward HIE models may refrain them from sharing their personal information with HIE networks. If patients are not willing to share their personal health information, incomplete, outdated, or inaccurate patient information will be stored in shared records of HIEs [13]. Accordingly, HIE efforts will fail in providing health care providers with reliable, useful, and integrated health information. Previous studies highlight that to maximize the full value of HIEs, it is important to evaluate patients' beliefs and perceptions about the widespread implementation of HIE networks [14]. Thus, public support is necessary for the long-term success and sustainability of HIE initiatives [8].

Different HIE models have attempted to clarify the process of electronic data sharing among health care entities. However, previous studies report that the general public is not completely aware of how health information is shared and used through the mainstream exchange mechanisms [15]. A number of studies highlight the importance of patient privacy and security concerns in the context of HIE implementation [16]. Patient concerns in medical practices include the volume of medical records collected and stored in health care organizations' databases, the possibility of privacy violations (eg, unauthorized access or hacked personal data), secondary use of medical records (eg, datamining purposes), lack of control over data collection practices, lack of transparency associated with sharing efforts, and lack of visibility about how such information will be used [17]. Patients will hold a positive attitude toward HIE networks when their health records are collected, stored, and exchanged confidentially [18]. According to Wright et al [19], if a patient's privacy and security needs are not met, he or she will become more likely to hide further health information from health care providers. Previous research indicates that patient decision to support HIE projects is a function of multiple factors such as type of information exchanged, privacy and security protections, and purpose behind the exchange [20]. Favorable attitude toward a HIE system is a result of a solid match between the HIE

XSL•FO
RenderX

mechanisms and transparency, security, as well as privacy requirements [5].

Blockchain in Health Care

Recent studies propose that blockchain is able to disrupt trusted business models mainly used in health care systems for information exchange purposes [21]. Considering the number of transactions (eg, information sharing) among health care entities and the expenses that hospitals experience in maintaining the HIE systems, the underlying blockchain technology of democratically sustained public ledgers of the records opens new and challenging opportunities for the health care industry. Blockchain can create an electronic context in which business transactions (such as information-sharing initiatives) between parties are conducted via a distributed community rather than a central authority or a single entity. This might essentially affect the transparency of the system and the role each entity plays [22]. Blockchain can also facilitate information exchange and coordination among health care entities and help patients become independent in the sharing of their medical records with providers. The mainstream HIE servers, depending on scale, are principally controlled by large corporations or health care institutions. This centralized control may raise privacy and security concerns because of abuses of power, which may result in secondary use of medical data, unauthorized access, and hacker attacks. Alternatively, the blockchain technology may promote a number of capabilities such as decentralization, security, privacy, breach resistance, and speed of certain features of the internet's infrastructure.

A great deal of interest has been reflected by recent studies to analyze the effects of blockchain-distributed ledger technologies on health care practices, and most of them are conceptual research [23]. However, little quantitative work has been conducted to investigate the exposure of HIE to blockchain technology. Little is also known about patients' attitudes toward the implementation of blockchain-enabled HIE networks, and it is still not clear if patients (as one of the key HIE stakeholders) are likely to opt in to the applications of this technology in HIE initiatives. Thus, more research is required to explore the core value of blockchain technology in the health care industry from health care consumers. Our work is among the first attempts to study the possible use of blockchain-based models in HIE from patients' perspectives. The results of this research can extend the current understanding of blockchain technology by helping health care organizations, health care communities, and policy makers identify the potential benefits and risks of using this technology in health care practices. From a practical standpoint, this study can be useful for HIE policy makers to better examine the patients' attitude toward the use of blockchain in HIEs, how it should be leveraged, and how patients can be impacted.

Research Background

In this section, first the shortcomings and problems with traditional health exchanges are explained to better clarify the research gap. Then, we investigate blockchain-based HIE as a potential solution to the problems.

Trust Issues in the Health Information Exchange Context

Trust plays a significant role in situations where there is a distance between consumers and vendors, such as in internet-dependent contexts [24]. HIE networks share individuals' health information electronically with other care providers to improve care coordination and enhance patient safety. HIE initiatives utilize sharing mechanisms with which health information is mostly transmitted without a patient's close supervision and control. Thus, patient's trust in the HIE is the core in this setting where a great deal of security concerns and privacy risks may entail [5]. Trust in HIE can predict patients' reactions to the implementation of HIE models because patients need to feel assured that the HIE networks will not compromise personal health information or misuse sensitive medical records [14]. Therefore, patients should trust HIE systems before they make an opt-in decision or disclose their personal health information.

Individual trust in HIE models can be a function of reliance on competence and integrity of sharing mechanisms [25]. Trust in HIE competence specifies the extent to which patients rely on technologically competent performance of the HIE to effectively disseminate health information between a wide variety of health organizations. Moreover, trust in HIE integrity refers to the belief that the agreement between the patients and HIE is reliable and honest. The lack of trust in HIE is mainly because of the distance imposed between patients and the actual users (health care organizations), lack of direct interactions between patients and HIE models, centralized control exerted by health care organizations, and the unfamiliar mechanisms used in the HIE system to share medical records electronically [26]. These characteristics create a setting that is more intangible than the traditional sharing methods (such as fax or mail). The mentioned reasons may make patient trust more critical in the settings where the 3 exchange models (ie, direct, query-based, and patient-mediated exchange) are mainly used.

Privacy Concern and Privacy Policy

HIE initiatives are developed to provide interorganizational networks in which patients' medical records are shared with a number of health care entities that are geographically scattered. When a networked-based technology (eg, HIE systems) deals with sharing sensitive information (such as health records), it is very likely that it exacerbates privacy concerns. Information privacy concerns may influence the validity and completeness of HIEs' patient databases, which may result in wasteful investment, inaccurate treatments, erroneous care planning, and higher mortality rates [12]. To avoid such issues, HIE networks should assure patients that their medical records would be well protected during exchange transactions. Thus, privacy policies should be clearly presented by health care organizations to highlight how sensitive health information will be used inside/outside the organizations and what security means will be utilized to protect such data from unauthorized access and secondary use [27]. The risks of violated privacy, information misuse, or unauthorized disclosure highlight the importance of developing a transparent privacy statement before patient medical records are disclosed and shared.

XSL•FO
RenderX

Previous studies emphasize that patients are highly concerned about losing control over how the mainstream HIE systems handle their health information [28]. The concern is mostly because of a lack of transparency on the HIEs' information practices and privacy policies. Privacy policies should be comprehensive and transparent enough to address all principles mentioned in the Health Insurance Portability and Accountability Act (HIPAA) [16]. The notice principle articulates what health information is collected and exchanged, what the purpose of data exchange is, how such information will be used internally, and whether patient data will be disclosed to third parties. The choice principle delineates the consent process and permission requirements. This dimension provides the choice to patients to put limits on providers for the exchange of health data. It also provides patients with the options to disclose such records to other third-party entities (eg, voluntary data disclosure for research purposes). The access principle entails granting the right to patients to obtain, review, and amend their personal information to ensure data accuracy and completeness. The security principle implies the adoption of reasonable measures and technical security steps to protect health information from unauthorized access, improper use, loss, unapproved alteration, or unanticipated disclosure during data exchange processes. The retention principle clarifies the acceptable duration of keeping and analysis of shared health information by health care providers for health care purposes. This dimension articulates the reasonable steps to permanently delete shared personal data if it is no longer required for the consented purpose. Finally, the enforcement principle highlights self-regulation such as privacy seals to protect information privacy by informing the public whether the exchange procedures correspond to the legal requirements [29]. Thus, highly transparent principles of privacy policies are able to demonstrate how safe, reliable, and dependent HIE networks are to reduce patients' concerns for information privacy.

Blockchain-Based Health Information Exchange as an Alternative

New ways of conducting business and operating economic activities are emerging through blockchain technology. Using dynamic shared ledgers, blockchain is able to facilitate recording business transactions between parties involved. Moreover, based on a peer-to-peer network of nodes, blockchain can also remove the need for intermediaries' interactions and direct control by third parties in running a business. According to Crosby et al [30], the underlying features of blockchain make it to be considered as a disruptive technology that has potentials to fundamentally change current business models. Most studies in the blockchain domain have investigated cryptocurrency for its technical properties [31]. However, blockchain technology has broader and deeper applications beyond cybercurrencies and can be used for other purposes than financial transactions. As the interest in this technology has been rising, blockchain is attracting a great deal of attention and investment from numerous projects in different sections [32]. Blockchain technology is transforming several industries, such as banking, electronic governance, electronic commerce (e-commerce), legal contracts, automation, logistics, and health care [33]. Owing to its underpinning technology, one of the most conceivable applications of blockchain is in establishing coordination and managing communication between networked companies (such as hospitals). In a networked business model, all the involved companies are required to uninterruptedly communicate and constantly update their supply chain components to track the latest status of orders, processes, and transactions.

Blockchain may also contribute to other organizational initiatives such as information exchange across affiliated/unaffiliated health care entities (ie, all parties involved in the health care process, such as physicians, hospitals, and clinics). It has been proposed that blockchain-based sharing models, which use immutability and built-in autonomy features of the blockchain, are able to efficiently track records of access to sensitive medical data stored in the cloud [34]. According to Xia et al [35], health care organizations can take advantage of the access control framework that is based on blockchain to facilitate and expedite medical data sharing with other institutions. This technology provides secure cryptographic techniques to strongly control the access to patient medical records stored and processed on cloud platforms. Relying on the robust security platform, the system can detect and validate users that have access to sensitive medical records and keep track of all sharing activities.

The technology behind blockchain enables anonymous/pseudonymous actors in sharing initiatives, especially in a cross-border setting such as HIE. Blockchain can also resolve technical issues such as security and scalability as it operates based on a peer-to-peer network with no central authority, administrator, or a firm controlling the transactions. This decentralized network prevents a single point of failure and a security breach [36]. Moreover, cryptographic protocol used by blockchain technology provides communications security over a computer network. Using smart contracts embedded in blockchain technology, health care institutions can tap into automated execution of business interactions to notably decrease the need for majority of office operations in the sharing process.

XSL•FO
RenderX

Table 1. Descriptions of health information exchange models.

HIE[a] model	Description	Reference
Direct exchange	Point-to-point information exchange in which a physician is able to share medical information with a known recipient over a secure network	Williams et al [6]
Query-based exchange	A single data repository that enables health care providers to share patient medical data with a centralized data warehouse. It also allows health care organizations to search for the required health information	Campion et al [5]
Patient-mediated exchange	This HIE model gives patients the ability to aggregate and manage their health information on the internet. Thus, patients can help share information between providers to track and monitor their own health	Rudin et al [43]
Blockchain-enabled exchange	A decentralized and trustless HIE model in which each block contains an episode of care and each node operates independently while following the sharing protocols. This model synthesizes medical data from patient-centered management tools and the EHR[b] systems to provide access only to authorized stakeholders through secure transactions	Jiang et al [44]

[a]HIE: health information exchange.

[b]EHR: electronic health record.

Blockchain technology is considered as a trustless distributed ledger to collect, store, share, analyze, and validate medical data exchange among different stakeholders (such as health care organizations, providers, and patients) [37]. Therefore, one of the most promising applications of blockchain in the health care domain is in health data transmissions between patients, providers, hospitals, and relevant entities [38]. Blockchain technology has been suggested as an underpinning infrastructure for HIE to improve medical data storage, information exchange, and medical record management [39]. Recent studies also propose adoption of blockchain-based data-sharing networks to analyze secondary medical data for biomedical research purposes [21]. Another stream of research focuses on the use of blockchain to store patient-centered outcomes [40] and patient consent data [41]. Several companies, such as Deloitte [42], Accenture [41], and Guardtime [34], have initiated adoption of blockchain-based systems to store, manage, and exchange patient care. Therefore, consistent with previous research, blockchain technology is able to contribute to the health care industry and HIE efforts. In summary, the main characteristics of the 4 HIE models examined in this study are described in Table 1.

Methods

Experiment Design

We designed 16 scenarios to analyze health care consumers' perceptions about the potentials and risks associated with the implementation of 4 possible HIE models (ie, direct, query based, patient centered, and blockchain based) built upon different architectures. The architectures of the 4 HIE models are different based on 2 factors: (1) transparency of privacy policy and (2) sensitivity of health information. In this study, we defined 2 extremes to examine the transparency of privacy policy used by the HIE models: strong versus weak. Moreover, we divided health information that could be exchanged through the HIE models into 2 types: sensitive versus nonsensitive. Figure 1 illustrates the 16 scenarios resulting from 4 HIE models, 2 types of privacy policy, and 2 types of health information.

Each scenario pertains to a separate experiment. Therefore, we conducted 16 separate experiments. As a between-subject experiment is a better choice than a within-subject experiment for attitude formation [45], in this study, we used between-subject experiments in which participants are randomly exposed to only 1 experiment. The total minimum sample required is 100 per experiment considering alpha=.05 and power beta=.95. As there are 6 main outcome variables in this study with 30 measures, we used minimum 120 respondents per experiment to reduce possible sampling errors. Table 2 shows the experimental design used in this study.

Question Development

Each experiment included 8 sections: experiment scenario, health information privacy concerns, opt-in intention measures, trust in competency of HIE technology, trust in integrity of exchange transactions, willingness to share information, perceived benefits of HIE, and finally, demographics as well as technology experience questions. In the scenario section, a hypothetical situation was clearly described in which consumers were randomly exposed to a HIE model with particular characteristics. Each scenario envisions a situation in which a health care provider is explaining one of the exchange models defined in Table 2 and asking respondents to read the described privacy policy as well as type of health information that will be shared through the mentioned HIE model. For instance, in experiment 1, 128 respondents were randomly exposed to a direct exchange model with a strong privacy policy designed to exchange highly sensitive health information. To ensure that respondents completely understood the assigned treatments, we provided a detailed description of the given exchange technology and its features in terms of HIE model and architecture. We avoided any negative or positive connotations with the HIE models to resolve the possible bias that may arise from use of favorable/unfavorable terms. Then, subjects were asked to reflect their perceptions and opinions about the described exchange mechanism by answering a series of questions mainly developed according to previous research.

XSL•FO
RenderX

Figure 1. Sixteen scenarios.

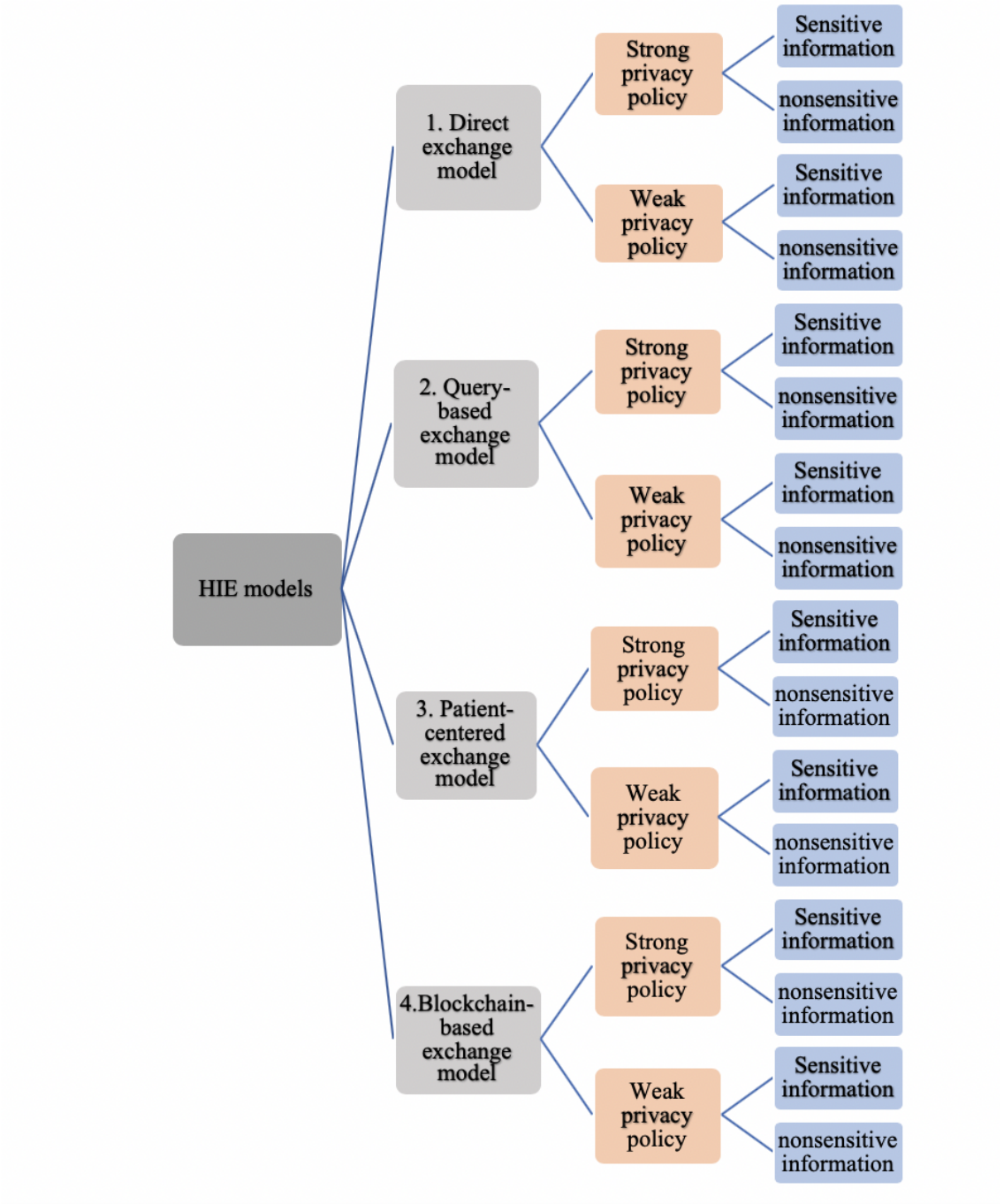


Table 2. Experimental design diagram.

HIE[a] model	HIE architecture							
	Strong policy/sensitive information		Strong policy/nonsensitive information		Weak policy/sensitive information		Weak policy/nonsensitive information	
	Experiment #	n	Experiment #	n	Experiment #	n	Experiment #	n
Direct exchange model	1	128	5	123	9	128	13	132
Query-based exchange model	2	128	6	131	10	124	14	126
Patient-centered exchange model	3	122	7	126	11	127	15	125
Blockchain-based exchange model	4	120	8	125	12	126	16	122

[a]HIE: health information exchange.

XSL•FO
RenderX

This study drew on the existing literature to measure the constructs included in the model, and minor changes were made to the instrument to fit the HIE context. To design the scenarios, we adapted the 6 dimensions of privacy policy transparency reported by Chua et al [29] and Wu et al [46] to distinguish between a strong and weak privacy policy. The sensitivity of health care information was categorized based on the classification of sensitive information provided by National Committee on Vital and Health Statistics [47]. Respondents' information privacy concern was measured based on their concern about the following items: collection, error, unauthorized access, and secondary use [48]. The scales used to measure trust in HIE technology's competency and trust in the exchange mechanism's integrity were adapted from a study conducted by Komiak and Benbasat [49]. Items measuring opt-in behavioral intention were adapted from previous research [50]. Items indicating willingness to disclose health information were adapted from the study by Zhang et al [51]. Items measuring perceived benefits were borrowed from factors suggested by previous studies [52,53]. All scales were measured on a 5-point Likert-type scale with 1 indicating *strongly disagree* and 5 indicating *strongly agree*. Finally, demographics and general technology experience questions were included at the end of the experiment (see Multimedia Appendix 1 for a description of the scenarios and questions).

We used the expert judgment approach to improve the content validity and completeness of our study. We sent the scenarios and questions to 5 professional health informatics practitioners and 3 blockchain experts. Then, the scenarios and questions were modified based on the experts' suggestions to ensure that they were clear and easy to understand for the public. Before conducting the main study, we also conducted a pilot test with 86 students at a large Southeastern university in the United States. We provided an open-ended essay box at the end of the survey for the students to comment on the clarity of the scenarios and the questions. Furthermore, we followed up on the comments by conducting interviews with the students to understand any ambiguity in the scenario and the surveys. We revised the scenario and the surveys based on the comments from the students before final data collection. To ensure the reliability and validity of the instrument, the Cronbach alpha was computed for each construct (privacy concern alpha=.85, trust in competency alpha=.76, trust in integrity alpha=.91, opt-in intention alpha=.88, willingness to share information alpha=.90, and perceived benefit alpha=.93). All the Cronbach alpha values were above the cutoff point of .7, which indicated that the instrument was internally consistent [54].

Recruitment and Participants

Data were collected in October 2018 using Amazon's Mechanical Turk (MTurk) to obtain a representative group of subjects. MTurk is used by a number of studies as an acceptable means to collect individual-level data [55,56]. Research in different domains (especially psychological and social behavior) recruits respondents through MTurk to analyze the perceptions of samples that are more representative of the general workforce, including a wide range of ages, ethnicities, and work experiences [57]. We defined a location filter to collect data from the United States. The 16 experiments were posed to MTurk at the same time. We used a randomizer function to assign respondents randomly to the 16 scenarios to minimize the likelihood that 1 respondent could participate in more than 1 experiment. Moreover, a microcode was activated in the survey to keep individuals from taking each experiment more than once. Finally, all experiments were also double-checked using generated respondent identification and internet protocol address to ensure that the respondents were unique between experiments. The incentive for participation was a monetary reward. The range of average completion time for the 16 experimental groups was between 21:49 and 32:36 min that implied acceptable responses in terms of timing.

The 16 experiments obtained data from 2013 respondents, ranging between 120 and 132 participants each. We matched the respondents across the 16 groups to avoid any potential problem of individual differences between groups. Results of chi-square tests show that there were no significant differences among participants in all 16 groups, and they are very similar in terms of the demographic variables (see Multimedia Appendix 2 for results of chi-square tests). For instance, the distribution of data related to gender (χ^2_{15}=12.1; *P*=.66), age (χ^2_{75}=92.8; *P*=.08), health status (χ^2_{60}=49.9; *P*=.91), household income (χ^2_{60}=59.1; *P*=.51), race (χ^2_{60}=81.5; *P*=.06), education level (χ^2_{75}=76.1; *P*=.44), employment status (χ^2_{60}=69.1; *P*=.19), and computer experience (χ^2_{60}=51.7; *P*=.77) was notably similar across the 16 scenarios. Thus, we had enough evidence to assume that matched groups were used in this study (see Multimedia Appendix 2 for respondent characteristics across the 16 experiments).

Results

Analysis of Variance Test

We used IBM SPSS Statistics 24 to perform analysis of variance (ANOVA) to examine whether the 16 groups are significantly different by our main outcome variables: privacy concerns, opt-in intention, trust in competency, trust in integrity, willingness to share information, and perceived benefits. Before performing ANOVA analysis, we ran the Levene test to examine the homogeneity of variance, as this is one of the fundamental assumptions of 1-way ANOVA. The results do not show enough evidence to hold the assumption of homogeneity of variance for outcome variables. Therefore, we conduct Welch ANOVA that presents the most power and lowest type I error rate when data violate the assumption of homogeneity of variances [58]. Table 3 shows the descriptive statistics (mean score, SE, and Welch values) and the significance of each outcome variable.

XSL•FO
RenderX

Table 3. Descriptives and summary of analysis of variance results.

Outcome variable and scenario #	Mean	SE	Welch	*P* value
Privacy concern			11.46	<.001
1	3.31	0.09		
2	3.51	0.09		
3	3.19	0.09		
4	3.05	0.10		
5	3.75	0.09		
6	3.73	0.08		
7	3.70	0.08		
8	3.71	0.08		
9	3.25	0.09		
10	3.34	0.08		
11	3.15	0.09		
12	2.98	0.10		
13	3.91	0.08		
14	3.84	0.08		
15	3.70	0.09		
16	3.60	0.09		
Trust in HIE[a] competency			5.64	<.001
1	3.46	0.07		
2	3.19	0.09		
3	3.40	0.07		
4	3.35	0.07		
5	3.16	0.09		
6	3.23	0.07		
7	3.09	0.07		
8	3.14	0.09		
9	3.38	0.08		
10	3.34	0.07		
11	3.32	0.08		
12	3.52	0.07		
13	3.00	0.09		
14	2.95	0.09		
15	3.00	0.09		
16	2.97	0.08		
Trust in HIE integrity			8.40	<.001
1	3.47	0.08		
2	3.10	0.09		
3	3.30	0.08		
4	3.29	0.08		
5	2.95	0.09		
6	3.04	0.08		
7	2.96	0.08		

XSL•FO
RenderX

Outcome variable and scenario #	Mean	SE	Welch	*P* value
8	2.93	0.09		
9	3.35	0.08		
10	3.36	0.07		
11	3.34	0.07		
12	3.52	0.07		
13	2.90	0.09		
14	2.91	0.08		
15	2.90	0.10		
16	2.85	0.09		
Opt-in intention			8.89	<.001
1	3.30	0.10		
2	2.89	0.11		
3	3.19	0.11		
4	3.08	0.11		
5	2.73	0.12		
6	2.63	0.10		
7	2.76	0.11		
8	2.62	0.11		
9	3.18	0.11		
10	3.14	0.10		
11	3.29	0.10		
12	3.50	0.09		
13	2.52	0.11		
14	2.58	0.11		
15	2.71	0.12		
16	2.65	0.11		
Willingness to share information			6.67	<.001
1	3.35	0.11		
2	2.95	0.11		
3	3.19	0.10		
4	3.19	0.11		
5	2.81	0.11		
6	2.70	0.10		
7	2.82	0.11		
8	2.68	0.11		
9	3.21	0.11		
10	3.13	0.10		
11	3.32	0.10		
12	3.36	0.10		
13	2.56	0.11		
14	2.70	0.11		
15	2.81	0.12		
16	2.73	0.11		

XSL•FO
RenderX

Outcome variable and scenario #	Mean	SE	Welch	*P* value
Perceived benefits			1.89	.02
1	3.84	0.08		
2	3.61	0.09		
3	3.68	0.08		
4	3.59	0.08		
5	3.59	0.09		
6	3.55	0.07		
7	3.59	0.07		
8	3.47	0.08		
9	3.76	0.08		
10	3.62	0.08		
11	3.75	0.09		
12	3.76	0.08		
13	3.52	0.08		
14	3.68	0.07		
15	3.63	0.08		
16	3.45	0.08		

[a]HIE: health information exchange.

The results of this table demonstrate that there are significant differences across different scenarios at the $P<.05$ level for the 6 outcome variables: privacy concern Welch (15, 752.9)=11.455, $P<.001$; trust in HIE competency Welch (15, 753)=5.64, $P<.001$; trust in integrity Welch (15, 753)=8.39, $P<.001$; opt-in intention Welch (15, 752.99)=8.89, $P<.001$; willingness to share information Welch (15, 753.07)=6.67, $P<.001$; and perceived benefits Welch (15, 752.9)=1.88, $P=.02$. Therefore, comparisons indicate that the levels of privacy concerns associated with sharing activities, trust in HIE models' competency, trust in integrity of sharing mechanisms, patients' opt-in intention to HIE initiatives, patients' willingness to disclose personal information, and perceived benefits of HIE networks significantly vary across the 4 HIE models, the 2 levels of health information sensitivity, and the 2 levels of privacy policy transparency. Furthermore, we conducted Games-Howell post hoc test, which is the multiple comparison procedure for means when variances and sample sizes are not equal, to identify which groups significantly differ from each other [59]. The following section describes the comparisons based on the 6 outcome variables used in this study.

Privacy Concern

We compared respondents' perception of privacy concerns associated with all 16 scenarios. For scenarios where a strong privacy policy is used to share sensitive health information, results reveal that privacy concern with blockchain technology is significantly lower than the direct exchange model ($t=-1.97$; $P=.03$) and the query-based model ($t=-3.49$; $P<.001$). Privacy concern for the patient-centered model is also significantly less than the query-based model ($t=-2.64$; $P<.001$). When comparing privacy concern between different HIE mechanisms for the scenarios where strong privacy policy is used to exchange nonsensitive information, we could not find any significant differences. In scenarios that use weak privacy policy for sharing sensitive information, we found that privacy concern with blockchain technology is significantly lower than the direct exchange model ($t=-2.82$; $P<.001$) and the query-based model ($t=-2.06$; $P=.02$). When respondents are exposed to scenarios with weak privacy concern for sharing nonsensitive health information, they express considerably lower privacy concern associated with blockchain technology compared with direct exchange ($t=-2.65$; $P<.001$) and query-based model ($t=-2.05$; $P=.02$). Overall, the results show that blockchain technology significantly reduces privacy concern among respondents compared with other HIE mechanisms regardless of the sensitivity of health information and strength of the privacy policy. Table 4 presents summary of significant results.

XSL•FO
RenderX

Table 4. Comparison of privacy concern across different health information exchange mechanisms.

Scenario	Health information exchange mechanism	*t*	*P* value
Strong policy/sensitive information			
Blockchain technology	Direct exchange model	−1.97	.03[a]
Blockchain technology	Query-based exchange model	−3.49	<.001[a]
Patient-centered exchange model	Query-based exchange model	−2.64	<.001[a]
Weak policy/sensitive information			
Blockchain technology	Direct exchange model	−2.82	<.001[a]
Blockchain technology	Query-based exchange model	−2.06	.02[a]
Weak policy/nonsensitive information			
Blockchain technology	Direct exchange model	−2.65	<.001[a]
Blockchain technology	Query-based exchange model	−2.05	.02[a]
Patient-centered exchange model	Query-based exchange model	−1.72	.04[a]

[a]The mean difference is significant at the .05 level.

Trust in Health Information Exchange Competency

Next, we compared the participants' responses to the level of trust in the capability of HIE mechanisms described in the sixteen 16 scenarios. According to Table 5, the results indicates that respondents who are exposed to strong privacy policies used to exchange sensitive information, express significantly more trust in the patient-centered exchange model (t=1.87; P=.03) and the direct exchange model (t=−2.39; P=.01) compared with the query-based model. We could not find significant differences in terms of trust in the competency of exchange technologies in other scenarios.

Trust in Exchange Integrity

Regarding respondents' level of trust in the integrity of the HIE mechanisms, the findings shown in Table 6 reveal that in the scenarios where sensitive information is shared with the help of strong privacy policies, there is a significant difference between blockchain versus query-based models (t=1.74; P=.04). In the same scenarios, our results show that trust in the integrity of the query-based model is significantly lower than that in the direct exchange model (t=−3.04; P=.001). There are no significant differences in terms of trust in the integrity and reliability of exchange mechanisms in other scenarios.

Opt-In Intention

Furthermore, we compared the intention of respondents to opt-in toward a HIE mechanism that was presented to them by the given scenarios. In scenarios where sensitive information was shared based on strong privacy policies, we found significant differences between the query-based model versus all other HIE mechanisms. Table 7 shows that the query-based model is found to be the least favorite model for respondents. When a weak privacy policy is used to share sensitive information, participants are significantly more inclined to opt-in toward the blockchain exchange model versus all other HIE mechanisms. Moreover, in scenarios where nonsensitive information is exchanged under weak privacy policies, the blockchain technology is more favorable compared with the direct (t=2.57; P=.005) and query-based models (t=2.22; P=.01).

Table 5. Comparison of trust in health information exchange competency across different health information exchange mechanisms.

Scenario (strong policy/sensitive information)	Health information exchange mechanism	*t*	*P* value
Direct exchange model	Query-based exchange model	−2.39	.01[a]
Patient-centered exchange model	Query-based exchange model	1.87	.03[a]

[a]The mean difference is significant at the .05 level.

Table 6. Comparison of trust in exchange integrity across different health information exchange mechanisms.

Scenario (strong policy/sensitive information)	Health information exchange mechanism	*t*	*P* value
Blockchain technology	Query-based exchange model	1.74	.04[a]
Query-based exchange model	Direct exchange model	−3.04	<.001[a]

[a]The mean difference is significant at the .05 level.

XSL•FO
RenderX

Table 7. Comparison of opt-in intention toward different health information exchange mechanisms.

Scenario	Health information exchange mechanism	*t*	*P* value
Strong policy/sensitive information			
Direct exchange model	Query-based exchange model	1.62	.04[a]
Blockchain Technology	Query-based exchange model	2.71	<.001[a]
Patient- centered exchange model	Query-based exchange model	1.93	.03[a]
Weak policy/sensitive information			
Blockchain technology	Direct exchange model	2.95	.001[a]
Blockchain technology	Query-based exchange model	2.63	.004[a]
Blockchain Technology	Patient-centered exchange model	1.70	.04[a]
Weak policy/nonsensitive information			
Blockchain technology	Direct exchange model	2.57	.005[a]
Blockchain technology	Query-based exchange model	2.22	.01[a]

[a]The mean difference is significant at the .05 level.

Willingness to Share Health Information

We further investigated whether respondents are willing to share their health information given the scenarios. In scenarios where sensitive information is shared under strong privacy policies, participants prefer blockchain technology significantly more than the query-based model (t=3.03; P=.001). Table 8 also shows that in the same scenarios, respondents express more willingness to share their information through the patient-centered model the than query-based model (t=2.01; P=.02). In scenarios where sensitive information is exchanged based on weak privacy policies, respondents exhibit significantly more willingness to share health information through blockchain technology compared with the direct model (t=5.07; P<.001), query-based model (t=5.61; P<.001), and patient-centered model (t=4.21; P<.001). In the same scenarios, we also found that respondents prefer the patient-centered model better than the query-based model (t=1.95; P=.03). Moreover, participants show more willingness toward blockchain technology for sharing nonsensitive information under weak privacy policies compared with the direct model (t=3.89; P<.001), query-based model (t=3.27; P<.001) and patient-centered model (t=2.001; P=.02). In the same scenarios, respondents also prefer the patient-centered model versus the direct model (t=2.09; P=.02) and the query-based-model (t=2.001; P=.02).

Perceived Benefits

With regard to the perceived benefits of HIE, there are no significant differences between the 4 HIE mechanisms given the different types of privacy policy and information sensitivity. Figures 2 to 7 display the differences in the means of different scenarios for each outcome variable.

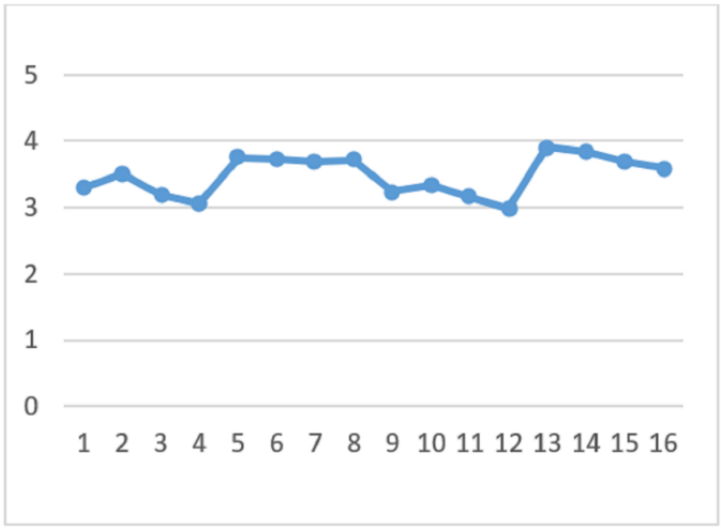
XSL•FO
RenderX

Table 8. Comparison of willingness to share information across different health information exchange mechanisms.

Scenario	Health information exchange mechanism	*t*	*P* value
Strong policy/sensitive information			
Blockchain technology	Query-based exchange model	3.03	.001[a]
Patient-centered exchange model	Query-based exchange model	2.01	.02[a]
Weak policy/sensitive information			
Blockchain technology	Direct exchange model	5.07	<.001[a]
Blockchain technology	Query-based exchange model	5.61	<.001[a]
Blockchain technology	Patient-centered exchange model	4.21	<.001[a]
Patient-centered exchange model	Query-based exchange model	1.95	.03[a]
Weak policy/nonsensitive information			
Blockchain technology	Direct exchange model	3.89	<.001[a]
Blockchain technology	Query-based exchange model	3.27	<.001[a]
Blockchain technology	Patient-centered exchange model	2.001	.02[a]
Patient-centered exchange model	Query-based exchange model	2.04	.02[a]
Patient-centered exchange model	Direct exchange model	2.09	.02[a]

[a]The mean difference is significant at the .05 level.

Figure 2. Differences in means of privacy concern.



XSL•FO
RenderX

Figure 3. Differences in means of trust in health information exchange competency.

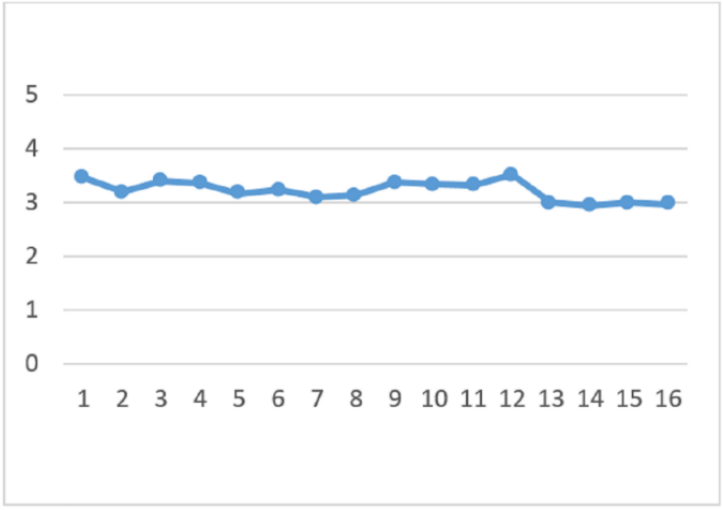


Figure 4. Differences in means of trust in exchange integrity.

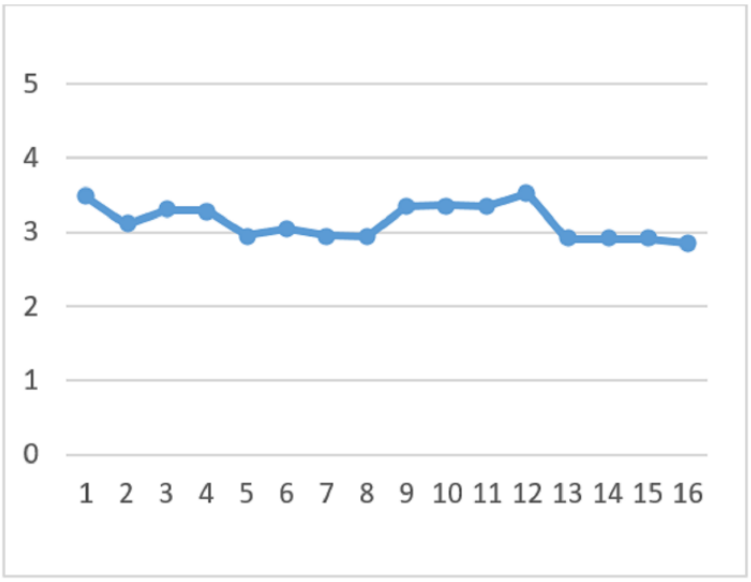
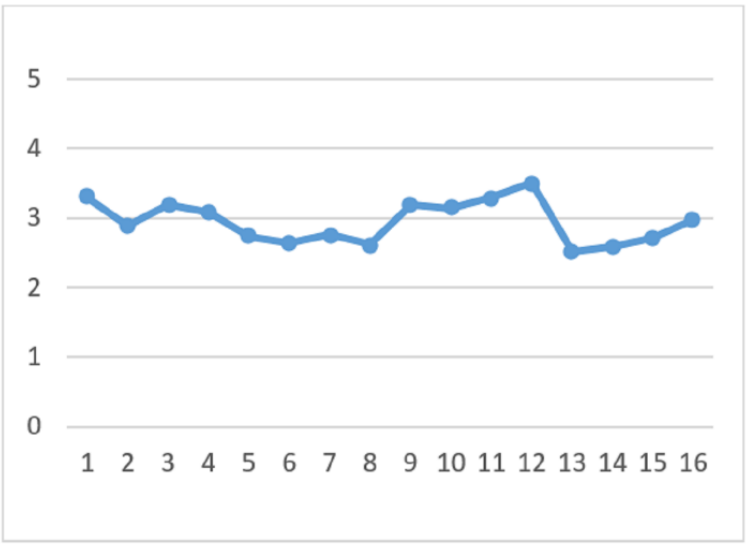


Figure 5. Differences in means of opt-in intention.



XSL•FO
RenderX

Figure 6. Differences in means of willingness to share information.

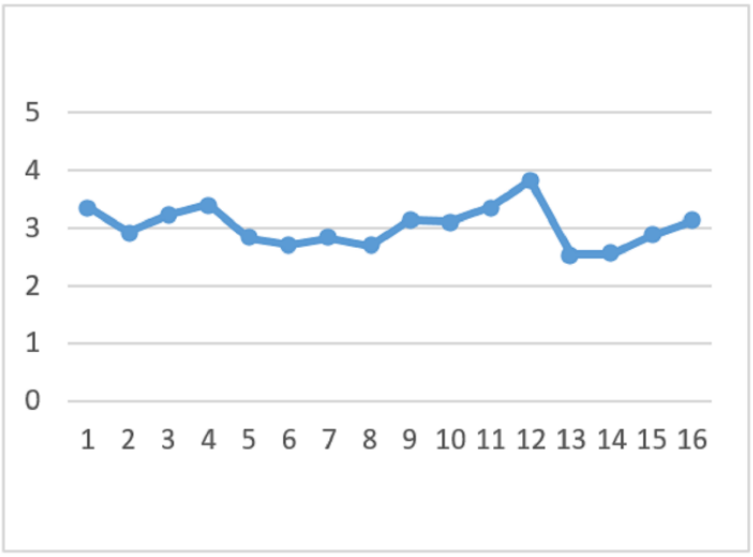
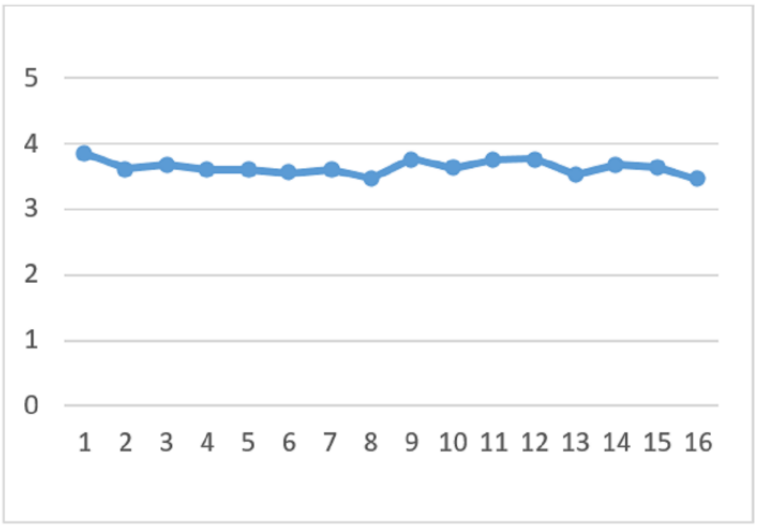


Figure 7. Differences in means of perceived benefits.



Discussion

Theoretical Implications

This study has implications for researchers conducting studies in the HIE context. Our study is different from previous research by examining patients' perspectives of 4 HIE models. This research is mainly designed to address how different models of HIE can affect patients' attitude toward electronic data exchange between health care providers. To do so, we investigated whether levels of patients' privacy concerns, perceived benefit of HIE, trust in HIE competency, trust in HIE integrity, willingness to share personal information, and opt-in intentions are different across multiple HIE models (ie, direct, query based, patient centered, and blockchain based). This study also contributes to the literature by providing new insights on how blockchain technology can be leveraged in the context of HIE and how patients may be affected.

Blockchain Technology

The content of the blockchain is information; thus, its use is adaptable in different business sectors. In line with the study by Liu and Tsyvinski [60], industries have different reactions to blockchain as they may benefit or become disrupted by this technology. Although there are several attempts among information systems (IS) scholars to recognize the applications of blockchain technology in different business contexts, significant theoretical contributions are still scarce, especially in the health care context. Economists, computer scientists, finance scholars, and IS scholars analyze blockchain technology and its adoption from various lenses. According to the existing blockchain literature, the majority of studies focus on its applications in financial transactions. This research domain can be divided into 5 main categories. One group of studies has focused on the potential use of cryptocurrency for illegal activities and has not examined the motives of mainstream users to adopt it for legitimate uses such as for e-commerce, information exchange, or money transfer [61]. The second group of studies has investigated cryptocurrency for its technical properties such as design science, cryptography, proof-of-work

XSL•FO
RenderX

algorithm, or exchange rates perspectives [31]. The third category of research has called readers' attention to analyze the differences between the technical, usability, and social characteristics of different forms of cryptocurrencies [62]. The fourth category of studies has attempted to distinguish cryptocurrency adopters from nonadopters based on either drivers or risks associated with its underlying technology [63]. The last category has used widely accepted adoption models (eg, technology acceptance model [TAM] and unified theory of acceptance and use of technology [UTAUT]) and mainly focused on the perceived usefulness and perceived ease of use in the context of cryptocurrency [64].

Blockchain-based sharing frameworks to facilitate the exchange of medical information between health care and research institutions are under development. These medical blockchain applications sufficiently control access to medical data stored and processed on cloud systems. They also offer secure cryptographic techniques to identify and authenticate users who have access to medical data to keep track of all exchange transactions [34]. Thus, sharing data for telemedicine and medical consultations in remote areas becomes more efficient. One of the main factors affecting the widespread adoption of blockchain-based HIE is social acceptance of this exchange mechanism. A patient may seek medical treatments and care planning from different health care organizations and providers. In a situation where patient-provider interactions are growing, a technology may be required to facilitate communications and medical records exchange without a centralized authority but relying on a distributed public ledger of all data exchange transactions. However, previous research on how patients would react to medical information sharing through blockchain is still in its nascent stage. In this study, we analyzed 6 outcome variables that need to be considered to measure the success and effectiveness of HIE models from patients' views. This work is among the first studies to empirically examine the potential role of blockchain in the HIE context. The results imply the significance of blockchain-based technology for health care applications when compared with the mainstream HIE models (ie, direct, query based, and patient mediate exchange).

Privacy Concerns

With regard to the privacy concern, blockchain-enabled HIE models in different scenarios (when either sensitive or nonsensitive health care information is shared under a strong or weak privacy policy) to receive favorable evaluations from our respondents. This is consistent with the study by Abdulnabi et al [18], which indicates that more decentralized models that increase patients' control over their medical data and exchange transactions will be more feasible and applicable approaches for HIE efforts. According to Vest and Gamm [3], using a centralized data repository in HIE initiatives has heightened privacy and security concerns for patients and created control issues for health care organizations. Moreover, an HIE model that uses blockchain technology to exchange sensitive information even under a weak privacy policy has more information privacy advantages from patients' perspectives compared with all other exchange models. Therefore, the findings show that the public considers blockchain as a more secure exchange model to share highly sensitive medical records regardless of privacy policy transparency. However, it can be discussed that open transparency of information (especially for sensitive records) during transfer can be obscure for consumers and should be addressed. This is in line with previous studies that indicate that blockchain is characterized as a decentralized, distributed, immutable, and transparent technology that can be used as permission-less or permissioned networks [23]. In the permission-less blockchain networks, any users can involve and participate without being authorized, and in the permissioned networks, only authorized users or organizations can participate. Owing to the overall sensitivity of health care information, stricter policy guidelines, and high compliance requirements in the context of HIE, the permissioned blockchain-based network would be a more secure option to enable electronic exchange of medical data with providers participating in other settings. This point is also highlighted by previous research indicating that although current blockchain technology underlying cryptocurrency is not fully anonymous, transaction anonymization for legitimate purposes (such as health care services) is desirable [65]. For example, in HIE networks, confidential health information should be handled with optimum security protocols.

Trust in Health Information Exchange Models

Sharing sensitive health information through a technology that is used by health care providers requires a new lens for understanding patients' trust in HIE technology. Regarding the trust in HIE competency and exchange integrity, a blockchain model even with no strong privacy policy is found as the most trustable model than other exchange mechanisms for sharing highly sensitive information. Consistent with this result, the public may believe that blockchain HIE has the necessary characteristics, technological capability, and features to be relied upon, regardless of presenting a comprehensive and transparent privacy policy for transmitting sensitive medical records (such as genetic information, mental health information, sexual health diseases, substance abuse, and addiction). Thus, blockchain HIE may heighten patients' cognitive dependence on HIE integrity and competence and win the trust of patients to exchange sensitive health-related information. This is consistent with previous studies that blockchain can be used as a reliable technology to share both highly sensitive medical data and less sensitive information such as current health statues (eg, fitness, diet, diseases, and treatments) or past medical/health information (eg, list of vaccinations and medications used) [35].

Consistent with previous studies, in the process of forming trust in technology (as an impersonal entity), consumers' awareness of the unknown should be resolved [66]. Previous research indicates that the public awareness about HIE mechanisms, functions, integrity, and security safeguards needs to be raised [2,26]. For example, one area could be the differences between the open transparency of information in cryptocurrency and blockchain-based HIE. The transparency of information in cryptocurrency means that all nodes in the network have the right to access the whole information related to financial transactions. However, this feature is not desirable for transmitting highly sensitive health information. To implement blockchain exchange methods to share sensitive health data across providers, it is required to develop security features (eg,

XSL•FO
RenderX

confidentiality, availability, and integrity), which is considered as one of the main aspects of blockchain technology. A blockchain-based HIE system is a decentralized framework where all medical records are confidential and the availability of such information does not rely on any third parties (eg, hospitals or providers). Furthermore, data integrity can be ensured because this form of HIE uses a distributed file system where participants in exchange activities will keep copies of all files, including the shared health information. Moreover, they agree to share, change, and update medical data by permission requests and consent processes. Therefore, the rational expectations about the HIE's ability to fulfil its obligations (cognitive trust in competence) and the rational reasons associated with the reliability of the HIE principles (cognitive trust in integrity) can be increased through raised awareness about the use of various types of blockchain innovations such as smart contract applications and permissioned networks.

Opt-In Intentions to Health Information Exchanges and Willingness to Share Information

Patients are considered as one of the most important stakeholders of any HIE efforts as the widespread implementation of HIE projects will not be feasible without their positive beliefs and attitudes toward the exchange models, their opt-in intentions to HIE initiatives, and their willingness to share health information [67]. The existing theories of information technology (IT) adoption (such as TAM and UTAUT) focus on users' intention to accept and use a technology [68]. However, in the HIE context, patients are not the main users. Patients are the beneficiaries of HIE initiatives, but they are not the final users. The users are health care professionals (ie, physicians and nurses), and the decision to adopt HIE is made at the practice/hospital level. However, it is critical to study whether patients will choose to opt in to HIE systems or they will not support such initiatives by hiding their personal health information. The results show that participants are most likely to opt in to blockchain HIE as a reliable technology to be used by health care entities to disseminate highly sensitive information even in the absence of a strong privacy notice. This finding is consistent with previous research highlighting that patients are more favorably disposed toward decentralized HIE models versus centralized exchange systems [69]. Furthermore, respondents are most willing to disclose sensitive health information to health care organizations, with the knowledge that such information may be exposed to other providers through a blockchain-based HIE even when privacy policy is not completely transparent.

The results manifest that with the current blockchain technology, patients may not feel skeptical about relying on blockchain-enabled HIE to manage the exchange of their highly sensitive information among a wide range of providers. This finding also emphasizes the importance of raising patient awareness of how the consent process and permissioned HIE networks operate in practice. Moreover, more efforts are required to improve the legal image of blockchain technology in health care to enable at-scale interoperability for information exchange, patient tracking, identity assurance, as well as validation among health care institutions and between patients and their providers [41]. Our findings also propose possible direct relationships of trust in blockchain HIE with patients' opt-in intentions and their willingness to disclose health information. Thus, a high level of trust in blockchain competence and integrity may encourage patients to opt in to this technology and disclose their sensitive health information when visiting a physician participating in a blockchain-based HIE network.

Perceived Benefits of Health Information Exchange

Pertaining to the perceived benefits of HIE, there is no significant difference across the scenarios. The 4 HIE models, regardless of different architectures (privacy policy and data sensitivity), receive the same level of benefits from patients. This means that although information privacy concerns can cause significant differences, all the exchange models are perceived to deliver comparable values. Thus, the instances that privacy policy dimensions are not stated transparently or conditions that highly sensitive medical data are likely to be shared will not significantly affect the core values expected from the HIE models. This is in line with previous studies that multiple exchange mechanisms may be used to fulfil different health care needs but the main purpose of all HIE models is to support care coordination, reduce health care costs, and improve patient safety [70]. Thus, patients may believe that regardless of what exchange model will be used for sharing information, HIE initiatives are generally able to improve communication among health care providers, reduce delays in care delivery, and advance quality of care planning.

Practical Implications

Patient Awareness About Blockchain-Based Health Information Exchange

There are also a number of important practical implications derived from this study. First, the findings suggest the importance of educating consumers about the use of blockchain technology in HIE mechanisms and sharing procedures. For instance, national educational programs, health conferences, and webinars that are easily accessible to a wide range of people can be administered to clearly publicize the key goals and advantages of blockchain-based HIE efforts. Educational forums available on official health websites, Web-based tutorials accessible on patient portals or Web-based health communities, and computerized help programs can be used by health care organizations to improve the transparency of blockchain applications in HIEs, broadcast their expected benefits, and increase public awareness and patient familiarity with this exchange mechanisms.

Second, regarding the importance of information privacy in blockchain-based HIE, health care providers should consider using tactics to increase the transparency and completeness of privacy policy and invest considerable effort in developing campaigns that leverage the power of blockchain image and reputation in health care. HIE policy makers should establish a broad marketing strategy to enhance patients' perceptions about the accountability and accuracy of privacy policies, which can foster patients' opt-in intention toward blockchain-enabled HIE services. Research implications suggest that HIE initiatives' managers should consider maximizing the transparency of

XSL•FO
RenderX

privacy policy dimensions to encourage consumers to read the privacy policy statements when data are subject to be exchanged through the blockchain networks.

Third, lack of public awareness about the blockchain-based HIE model as well as the components of its privacy statement may impede the progress of sharing information between providers because of the lack of patients' support for HIE. This study suggests that both physicians and health care organizations (such as hospitals) can directly play an important role in persuading patients to give consent to sharing medical records using blockchain-enabled HIE. Physicians' role may be more effective because they have face-to-face encounters with patients and during consultations, they can enlighten the patients about the benefits of using blockchain in HIEs, and how they could be impacted. For instance, health care professionals can explain how HIE, which is enabled by blockchain can help physicians detect diseases faster, coordinate treatments with other providers, and finally, improve patient safety. Hospitals can also influence how patients shape opt-in decisions toward blockchain-based HIE by educating them using brochures, leaflets, diagrams, and fact sheets that are comprehensible for an average person. These efforts should be able to clearly highlight why health information is shared, what types of information can be exchanged, how such information is shared from one point to another, what exchange mechanisms are used, who can access the medical data, what security safeguards will protect their records, and how often the transmission takes place.

Potential Benefits of Blockchain-Based Health Information Exchange

Relying on the key findings as well as characteristics and features of blockchain, the main benefits of using blockchain for improving medical record sharing among health care organizations are discussed in the following section. Decentralized management of the blockchain technology can notably contribute to HIE by providing patient-managed health care records. In these platforms, patients are considered as the owner of their medical records and are able to efficiently control access to such information [37]. This aspect can also help patients reduce all possible barriers associated with obtaining copies of their medical information and potential risks related to sharing them with other health care organizations. In a blockchain-based HIE, each block can contain an episode of care and each node operates independently while following the sharing protocols. Blockchain has the potential to become an electronic health information pool by synthesizing medical data from patient-centered management tools and the EHR systems to provide access only to authorized stakeholders (such as patients and providers). The peer-to-peer architecture of blockchain also enables health care institutions to keep control of their own IT resources and collaborate with other organizations to enhance information sharing initiatives without ceding control [71]. Thus, incorporating blockchain into HIE is appropriate for health care providers/organizations that seek to cooperate with each other with no centralized management intermediary. On the contrary, most of HIE mechanisms (eg, direct model or lookup networks) are centrally managed.

The immutable audit trail is another characteristic of blockchain technology that is likely to contribute to HIE. On the basis of this aspect, patient health information is not changeable in any steps of the data-sharing initiatives. Thus, the medical records that are stored in the private blockchain cloud cannot be altered, manipulated, or removed by any entities participating in HIE initiatives such as health care providers and organizations [37]. Furthermore, patient medical records that are generated and shared with health care providers through a blockchain-based sharing platform are trackable and timestamped.

Managing patient consent records during data-sharing processes can be improved by the data provenance of blockchain technology. This aspect can help the owner of medical records to change the ownership or give permission to other entities to view, process, and share such information using the cryptographic protocols. Moreover, patients or providers can trace the source of data and verify legitimacy as well as accuracy of records to be used for exchange purposes. Thus, using blockchain-based HIE, the source of medical records is detected and any ownership transfer in each block will be transparent and available to everyone involved in the data-sharing efforts.

Blockchain is built upon distributed technology that does not suffer from a single point of failure. Relying on this feature, patient health information can be collected, stored, and shared on a decentralized network, where there is no central institution that could be hacked or compromised. This robustness feature has the potential to decrease the risk of patient recordkeeping as medical data cannot be faked or manipulated. Moreover, one of the main threats related to the mainstream HIE models is unavailability of patient data when incomplete or inaccurate patient information is stored in shared records [12]. This issue can be resolved by blockchain technology as each node in the network has a copy of historical medical records and is able to continuously update such data. This characteristic may guarantee that the electronic medical records of patients are continuously available in real time [72]. Real-time access to patient data is one of the main promises of HIE efforts that enable providers to advance care coordination, detect epidemics rapidly, and improve care delivery in emergency situations [73].

All information exchange initiatives in the United States health care industry (such as HIE projects) fall under HIPAA security rules [25]. Under HIPAA, security policies and procedures should be implemented to prevent, detect, and correct security violations [74]. For example, a thorough analysis of the potential risks and vulnerabilities to the confidentiality, integrity, and availability of electronic health information held by the covered entity should be conducted before exchanging any information. Moreover, procedures for the authorization and supervision of members who use electronic protected health information should be clarified [16]. For instance, security procedures should determine whether the access of a health care organization to electronic health information is appropriate or should be terminated. On the basis of HIPAA guidelines, procedures are implemented to verify that a physician or entity seeking access to electronic health information is the one claimed [75]. For instance, technical security measures and encryption mechanisms are implemented to guard against unauthorized

XSL•FO
RenderX

access to electronic health information that is being transmitted over an electronic communications network (such as HIE).

A number of studies have argued that privacy and security concerns are identified as the most pressing barriers to widespread consumer participation in the implementation of mainstream HIE models [53]. Privacy policies of HIE efforts should be comprehensive and transparent enough to address all the principles mentioned in HIPAA [16]. One of the main advantages of blockchain technology that can be utilized by HIE models is improving safety, integrity, and confidentiality of patient health information using cryptographic algorithms and consent recording systems. Episodes of medical care can be stored in blocks and only decrypted for exchange purposes with the patient's private key. Even if the distributed network is breached by a malicious entity, with current technology means, it is extremely unlikely that patient data can be illegally accessed by unauthorized parties. Therefore, blockchain-enabled sharing platforms have the potential to connect a vast number of patients, health care providers, and health care organizations to exchange variety of medical records while information privacy and security are protected.

Plausible Challenges of Blockchain-Based Health Information Exchange

The objective of this study was not to propose blockchain as the most advantageous method of information exchange in the health care industry. Results of our research indicate that there are still a number of criticisms attributed to blockchain-based solutions in the health care area. In this section, the main shortcomings of blockchain are highlighted to imply that current blockchain solutions need some necessary modifications to be implemented in the HIE context.

Although blockchain is an appropriate means to facilitate interoperability, current studies have also emphasized that the open transparency of data during exchange transactions is not desirable in health care applications [41]. In the HIE context, identifiable information of patients is highly sensitive. The key objective of HIPAA compliance is that information exchange must be protected against a confidentiality breach. The end-to-end workflow of a blockchain-based HIE (ie, entering, processing, and delivering of health data) must be HIPAA compliant. Any personal health information accessed by the blockchain-enabled HIE must be encrypted and securely managed by parties interacting with this HIE model. A blockchain-based HIE should not make all personal information publicly available so it can securely store and manage sensitive data. Blockchain-enabled HIE should ensure the anonymity of each identity and transaction using unique authentication protocol (data protection methods such as tokenization or masking) [76]. Thus, each data exchange performed by a user should not be linked to the user and the ownership of the key should remain anonymous. Moreover, privacy policies designed for blockchain-based HIE can provide different levels of data access and, if required, time-limited access. Another way to alleviate the open transparency issue is encrypting sensitive health records on the network of blockchain-based HIE [77]. Recent studies also propose that sensitive medical records can be stored off-blockchain network and only encrypted links and permission information should be exchanged on network [38]. According to Ekblaw et al [78], data exchange protocols can be automated using smart contracts to attenuate this risk.

Another potential challenge with the adoption of blockchain in HIE networks is the speed of transactions. Depending on the authentication and verification protocol used in blockchain, data exchange processes could be time consuming. This could challenge the real-time communication, coordination, and data sharing among health care providers, which is critical in many health care situations [23]. According to Linn and Koo [72], ongoing verified exchange transactions can only be stored in blocks instead of the complete past medical histories. Another plausible solution is to implement blockchain-based platforms that provide higher transaction speed compared with the Bitcoin network [79].

The risk of a 51% attack has been considered as an important threat to blockchain networks [80]. This attack, which occurs when the whole network is controlled by attackers or malicious nodes, could critically threaten the security of HIE platforms. HIE can adopt permissioned blockchain networks in which malicious nodes are not able to randomly contribute to the network, and in turn, the risk of a 51% attack could be minimized. For instance, implementation of a virtual private network in which medical records are stored and exchanged on private cloud resources complied with HIPAA can notably mitigate this risk [81].

Finally, it should be mentioned that the spread of the blockchain-based framework in health care practices might be challenging, particularly in developing countries that do not have adequate technical infrastructure and social support. Moreover, the long-term success of blockchain-based HIE needs favorable attitude and active participation of all stakeholders (such as physicians, health care organizations, and patients). According to Dixon et al [82], HIE projects may become ineffective and disabled because of a lack of participation and support from HIE stakeholders. With respect to patients, there is a need to increase public awareness about blockchain technology. For instance, national educational programs such as educational videos or webinars can be used by health care providers to convey key information about blockchain-based HIE and how it facilitates the sharing of medical data securely with and between health care providers. Patients should be educated on the aspects of blockchain to realize how the technology is able to exchange sensitive medical data securely, improve confidentiality of all sharing activities, enable patients to track who can access episode-of-care data, and increase patient control over their medical records.

Limitations and Future Research

Similar to other studies, our research has some limitations that call for additional work. We began this study by reflecting on patients' perceptions about the implementation of 4 HIE models. Researchers coming from a different starting point could contribute to this research stream in different ways. We raise this point, not to defend our view or to deflect criticism, but simply to clarify the scope of our paper and motivate future research that takes different perspectives or assumptions. For instance, future work can examine health care professionals'

XSL•FO
RenderX

perspectives or investigate health care organizations' requirements and limitations on the implementation of blockchain-based HIE alternatives. This study is mainly designed based on the hypothetical scenarios that clearly define the use of 4 HIE models under different circumstances (ie, privacy policy and type of information). Relying on existing literature, expert judgment approach, and pilot testing, we provided clear definitions by articulating the HIE models, privacy policy, and data sensitivity to reduce possible ambiguity. However, as HIE still is a relatively new technology, there was a small chance that some respondents did not comprehend the scenarios completely. Thus, we suggest that further studies use samples who have experience with the HIE models.

Consistent with the results of this study, further research can also develop and empirically test a causal model using the outcome variables proposed by this study to predict the success of blockchain in HIE initiatives from consumers, health care professionals, and hospital managers' perspectives. Health care industry is considered as a highly regulated environment. Future studies can extend this work by identifying approaches to address governance conflicts arising from the technology being used in the health care context. It can also be of interest for future research to investigate the role of regulatory bodies in keeping control, on the one hand, and having systems that run on their own, on the other. In this study, we discussed the key risks involved along with several plausible solutions related to the adoption of blockchain technology in the HIE context. Future research is required to shed more light on the design and implementation of blockchain-enabled HIE applications. Finally, this study provides a footstone for further theoretical development and practical investigation. For instance, future work can study the return on investment and cost impact to health care delivery as a result of a blockchain-enabled HIE implementation. Moreover, the legal and policy implications/requirements can be addressed by further research.

Conclusions

Blockchain is considered as one of the most important technologies that can be applied in many sectors in the future. One of the most interesting cases of blockchain technology application is in health care domains. Research on the use of blockchain technology in the health care context is still in its early stages, and its widespread adoption needs further efforts. This work uses an experimental approach to better articulate the prospective application of blockchain technology in creating an infrastructure for sharing medical records. The findings indicate that blockchain technology has a great potential to be integrated in existing HIE architectures to improve system transparency, patient consent tracking, and privacy protection of information exchange initiatives. Blockchain-based HIE is able to provide a platform for data exchange that does not need a centralized authority to operate. This aspect promotes a protocol supporting a network-based communication between patients and physicians and a well-organized coordination among health care organizations to accurately diagnose diseases, provide timely treatments, and improve patient safety. According to the results of this study, patients perceive that blockchain technology can be a reliable replacement for current exchange models, which are mainly managed by mainstream bureaucratic systems or large institutions with centralized control (such as hospitals). Consistent with results, we also discuss the key benefits and possible risks of adopting blockchain technology in HIE efforts. This research can serve as a foundation for future studies in the domain of blockchain-based HIE.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Experiments: scenarios and questions.

[DOCX File, 26KB - jmir_v21i6e14184_app1.docx]

Multimedia Appendix 2

Respondent characteristics across the 16 experiments.

[DOCX File, 30KB - jmir_v21i6e14184_app2.docx]

References

1. Cherry JC, Dryden K, Kobb R, Hilsen P, Nedd N. Opening a window of opportunity through technology and coordination: a multisite case study. Telemed J E Health 2003;9(3):265-271. [doi: 10.1089/153056203322502650] [Medline: 14611694]
2. Edaibat EA, Dever J, Stuban SM. System dynamics simulation modeling of health information exchange (HIE) adoption and policy intervention: a case study in the State of Maryland. Oper Res Health Care 2017 Mar;12:60-70. [doi: 10.1016/j.orhc.2017.02.001]
3. Vest JR, Gamm LD. Health information exchange: persistent challenges and new strategies. J Am Med Inform Assoc 2010;17(3):288-294 [FREE Full text] [doi: 10.1136/jamia.2010.003673] [Medline: 20442146]
4. Lenert L, Sundwall D, Lenert ME. Shifts in the architecture of the Nationwide Health Information Network. J Am Med Inform Assoc 2012;19(4):498-502 [FREE Full text] [doi: 10.1136/amiajnl-2011-000442] [Medline: 22268218]

XSL•FO
RenderX

5. Campion TR, Edwards AM, Johnson SB, Kaushal R, HITEC Investigators. Health information exchange system usage patterns in three communities: practice sites, users, patients, and data. Int J Med Inform 2013 Sep;82(9):810-820. [doi: 10.1016/j.ijmedinf.2013.05.001] [Medline: 23743323]
6. Williams C, Mostashari F, Mertz K, Hogin E, Atwal P. From the Office of the National Coordinator: the strategy for advancing the exchange of health information. Health Aff (Millwood) 2012 Mar;31(3):527-536. [doi: 10.1377/hlthaff.2011.1314] [Medline: 22392663]
7. Walker J, Pan E, Johnston D, Adler-Milstein J, Bates DW, Middleton B. The value of health care information exchange and interoperability. Health Aff (Millwood) 2005;Suppl Web Exclusives:W5-10. [doi: 10.1377/hlthaff.w5.10] [Medline: 15659453]
8. Ancker JS, Edwards AM, Miller MC, Kaushal R. Consumer perceptions of electronic health information exchange. Am J Prev Med 2012 Jul;43(1):76-80. [doi: 10.1016/j.amepre.2012.02.027] [Medline: 22704751]
9. Rudin R, Volk L, Simon S, Bates D. What affects clinicians' usage of health information exchange? Appl Clin Inform 2011 Jan 1;2(3):250-262 [FREE Full text] [doi: 10.4338/ACI-2011-03-RA-0021] [Medline: 22180762]
10. Patel VN, Abramson E, Edwards AM, Cheung MA, Dhopeshwarkar RV, Kaushal R. Consumer attitudes toward personal health records in a beacon community. Am J Manag Care 2011 Apr;17(4):e104-e120 [FREE Full text] [Medline: 21774099]
11. Esmaeilzadeh P. The effects of public concern for information privacy on the adoption of Health Information Exchanges (HIEs) by healthcare entities. Health Commun 2018 May 8:1-10. [doi: 10.1080/10410236.2018.1471336] [Medline: 29737872]
12. Whiddett R, Hunter I, Engelbrecht J, Handy J. Patients' attitudes towards sharing their health information. Int J Med Inform 2006 Jul;75(7):530-541. [doi: 10.1016/j.ijmedinf.2005.08.009] [Medline: 16198142]
13. Dimitropoulos L, Patel V, Scheffler SA, Posnack S. Public attitudes toward health information exchange: perceived benefits and concerns. Am J Manag Care 2011 Dec;17(12 Spec No):SP111-SP116 [FREE Full text] [Medline: 22216769]
14. Tripathi M, Delano D, Lund B, Rudolph L. Engaging patients for health information exchange. Health Aff (Millwood) 2009;28(2):435-443. [doi: 10.1377/hlthaff.28.2.435] [Medline: 19276000]
15. Wang JY, Ho HY, Chen JD, Chai S, Tai CJ, Chen YF. Attitudes toward inter-hospital electronic patient record exchange: discrepancies among physicians, medical record staff, and patients. BMC Health Serv Res 2015 Jul 12;15:264 [FREE Full text] [doi: 10.1186/s12913-015-0896-y] [Medline: 26162748]
16. Dimitropoulos L, Rizk S. A state-based approach to privacy and security for interoperable health information exchange. Health Aff (Millwood) 2009;28(2):428-434. [doi: 10.1377/hlthaff.28.2.428] [Medline: 19275999]
17. Agaku IT, Adisa AO, Ayo-Yusuf OA, Connolly GN. Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. J Am Med Inform Assoc 2014;21(2):374-378 [FREE Full text] [doi: 10.1136/amiajnl-2013-002079] [Medline: 23975624]
18. Abdulnabi M, Al-Haiqi A, Kiah ML, Zaidan AA, Zaidan BB, Hussain M. A distributed framework for health information exchange using smartphone technologies. J Biomed Inform 2017 Dec;69:230-250 [FREE Full text] [doi: 10.1016/j.jbi.2017.04.013] [Medline: 28433825]
19. Wright A, Soran C, Jenter CA, Volk LA, Bates DW, Simon SR. Physician attitudes toward health information exchange: results of a statewide survey. J Am Med Inform Assoc 2010;17(1):66-70 [FREE Full text] [doi: 10.1197/jamia.M3241] [Medline: 20064804]
20. Willison DJ, Schwartz L, Abelson J, Charles C, Swinton M, Northrup D, et al. Alternatives to project-specific consent for access to personal information for health research: what is the opinion of the Canadian public? J Am Med Inform Assoc 2007;14(6):706-712 [FREE Full text] [doi: 10.1197/jamia.M2457] [Medline: 17712084]
21. Peterson K, Deeduvanu R, Kanjamala P, Boles K. HealthIT.gov. 2016. A Blockchain-Based Approach to Health Information Exchange Networks URL: https://www.healthit.gov/sites/default/files/12-55-blockchain-based-approach-final.pdf [accessed 2019-06-03] [WebCite Cache ID 78rXAm46q]
22. Chen Y. Blockchain tokens and the potential democratization of entrepreneurship and innovation. Bus Horiz 2018 Jul;61(4):567-575. [doi: 10.1016/j.bushor.2018.03.006]
23. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 1;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
24. Baptista G, Oliveira T. Understanding mobile banking: the unified theory of acceptance and use of technology combined with cultural moderators. Comput Hum Behav 2015 Sep;50:418-430. [doi: 10.1016/j.chb.2015.04.024]
25. McGraw D, Dempsey JX, Harris L, Goldman J. Privacy as an enabler, not an impediment: building trust into health information exchange. Health Aff (Millwood) 2009;28(2):416-427. [doi: 10.1377/hlthaff.28.2.416] [Medline: 19275998]
26. Eden KB, Totten AM, Kassakian SZ, Gorman PN, McDonagh MS, Devine B, et al. Barriers and facilitators to exchanging health information: a systematic review. Int J Med Inform 2016 Apr;88:44-51 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.01.004] [Medline: 26878761]
27. Li Y, Stweart W, Zhu J, Ni A. Online Privacy Policy of the Thirty Dow Jones Corporations: Compliance with FTC Fair Information Practice Principles and Readability Assessment. San Bernardino: California State University; 2012:65-89.

XSL•FO
RenderX

28. Kim KK, Joseph JG, Ohno-Machado L. Comparison of consumers' views on electronic data sharing for healthcare and research. J Am Med Inform Assoc 2015 Jul;22(4):821-830 [FREE Full text] [doi: 10.1093/jamia/ocv014] [Medline: 25829461]
29. Chua H, Herbland A, Wong S, Chang Y. Compliance to personal data protection principles: a study of how organizations frame privacy policy notices. Telemat Inf 2017 Jul;34(4):157-170. [doi: 10.1016/j.tele.2017.01.008]
30. Crosby M, Pradan P, Sanjeev V, Kalyanaraman V. Blockchain technology: beyond bitcoin. Appl Innov Rev 2016 Jun(2):6-19 [FREE Full text]
31. Li X, Wang C. The technology and economic determinants of cryptocurrency exchange rates: the case of bitcoin. Deci Supp Syst 2017 Mar;95:49-60. [doi: 10.1016/j.dss.2016.12.001]
32. Nofer M, Gomber P, Hinz O, Schiereck D. Blockchain. Bus Inf Syst Eng 2017 Mar 20;59(3):183-187. [doi: 10.1007/s12599-017-0467-3]
33. Underwood S. Blockchain beyond bitcoin. Commun ACM 2016 Oct 28;59(11):15-17. [doi: 10.1145/2994581]
34. Giungato P, Rana R, Tarabella A, Tricase C. Current trends in sustainability of bitcoins and related blockchain technology. Sustainability 2017 Nov 30;9(12):2214. [doi: 10.3390/su9122214]
35. Xia Q, Sifah E, Smahi A, Amofa S, Zhang X. BBDS: blockchain-based data sharing for electronic medical records in cloud environments. Information 2017 Apr 17;8(2):44. [doi: 10.3390/info8020044]
36. Fromknecht C, Velicanu D, Yakoubov S. A decentralized public key infrastructure with identity retention. Cryptol ePrint Arch 2014 Nov:1-16 [FREE Full text]
37. Yue X, Wang H, Jin D, Li M, Jiang W. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J Med Syst 2016 Oct;40(10):218. [doi: 10.1007/s10916-016-0574-6] [Medline: 27565509]
38. Azaria A, Ekblaw A, Vieira T, Lippman A. MedRec: Using Blockchain for Medical Data Access and Permission Management. 2016 Aug Presented at: 2016 2nd International Conference on Open and Big Data (OBD); August, 22-24, 2016; Vienna, Austria p. 25-30 URL: https://ieeexplore.ieee.org/abstract/document/7573685 [doi: 10.1109/OBD.2016.11]
39. Mettler M. Blockchain technology in healthcare: The revolution starts here. : IEEE; 2016 Presented at: 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom); September 14-16, 2016; Munich, Germany p. 1-3. [doi: 10.1109/HealthCom.2016.7749510]
40. Cichosz SL, Stausholm MN, Kronborg T, Vestergaard P, Hejlesen O. How to use blockchain for diabetes health care data and access management: an operational concept. J Diabetes Sci Technol 2019 Mar;13(2):248-253. [doi: 10.1177/1932296818790281] [Medline: 30047789]
41. Brodersen C, Kalis B, Leong C, Mitchell E, Pupo E, Truscott A. Semantic Scholar. 2016. Blockchain: Securing a New Health Interoperability Experience URL: https://pdfs.semanticscholar.org/8b24/dc9cffeca8cc276d3102f8ae17467c7343b0.pdf [accessed 2019-06-03] [WebCite Cache ID 78rbCxfED]
42. Krawiec RJ, Barr D, Killmeyer J, Filipova M, Nesbitt A, Israel A, et al. Semantic Scholar. UK: Deloitte Development LLC; 2016. Blockchain: Opportunities for Health Care URL: https://pdfs.semanticscholar.org/b992/77c3eecfe6d3dd784fe572a45780ffd040e2.pdf [accessed 2019-06-03] [WebCite Cache ID 78rbbsp6H]
43. Rudin R, Volk L, Simon S, Bates D. What affects clinicians' usage of health information exchange. Appl Clin Inform 2011 Jan 1;2(3):250-262 [FREE Full text] [doi: 10.4338/ACI-2011-03-RA-0021] [Medline: 22180762]
44. Jiang S, Cao J, Wu H, Yang Y, Ma M, He J. Blochie: a blockchain-based platform for healthcare information exchange. : IEEE; 2018 Presented at: 2018 IEEE International Conference on Smart Computing (SMARTCOMP); June 18-20, 2018; Taormina, Italy p. 49-56 URL: https://ieeexplore.ieee.org/abstract/document/8421331 [doi: 10.1109/SMARTCOMP.2018.00073]
45. Charness G, Gneezy U, Kuhn M. Experimental methods: between-subject and within-subject design. J Econ Behav Organ 2012 Jan;81(1):1-8. [doi: 10.1016/j.jebo.2011.08.009]
46. Wu K, Huang SY, Yen DC, Popova I. The effect of online privacy policy on consumer privacy concern and trust. Comput Hum Behav 2012 May;28(3):889-897. [doi: 10.1016/j.chb.2011.12.008]
47. Caine K, Hanania R. Patients want granular privacy control over health information in electronic medical records. J Am Med Inform Assoc 2013 Jan 1;20(1):7-15 [FREE Full text] [doi: 10.1136/amiajnl-2012-001023] [Medline: 23184192]
48. Van Slyke C, Shim JT, Johnson R, Jiang JJ. Concern for information privacy and online consumer purchasing. J Assoc Inf Sci 2006;7(6):16 [FREE Full text]
49. Komiak S, Benbasat I. The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS Q 2006;30(4):941-960 [FREE Full text] [doi: 10.2307/25148760]
50. Angst CM, Agarwal R. Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS Q 2009 Jun;33(2):339-370 [FREE Full text]
51. Zhang X, Liu S, Chen X, Wang L, Gao B, Zhu Q. Health information privacy concerns, antecedents, and information disclosure intention in online health communities. Inf Manage 2018 Jun;55(4):482-493. [doi: 10.1016/j.im.2017.11.003]
52. O'Donnell HC, Patel V, Kern LM, Barrón Y, Teixeira P, Dhopeshwarkar R, et al. Healthcare consumers' attitudes towards physician and personal use of health information exchange. J Gen Intern Med 2011 Sep;26(9):1019-1026 [FREE Full text] [doi: 10.1007/s11606-011-1733-6] [Medline: 21584839]

XSL•FO
RenderX

53. Patel VN, Dhopeshwarkar RV, Edwards A, Barrón Y, Sparenborg J, Kaushal R. Consumer support for health information exchange and personal health records: a regional health information organization survey. J Med Syst 2012 Jun;36(3):1043-1052. [doi: 10.1007/s10916-010-9566-0] [Medline: 20703633]
54. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. J Market Theory Pract 2014 Dec 8;19(2):139-152. [doi: 10.2753/MTP1069-6679190202]
55. Marge M, Banerjee S, Rudnicky A. Using the Amazon Mechanical Turk for transcription of spoken language. : IEEE; 2010 Presented at: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing; March 14-19, 2010; Dallas, TX, USA p. 5270-5273 URL: https://ieeexplore.ieee.org/abstract/document/5494979 [doi: 10.1109/ICASSP.2010.5494979]
56. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. Judgm Decis Mak 2010;5(5):411-419 [FREE Full text]
57. Behrend TS, Sharek DJ, Meade AW, Wiebe EN. The viability of crowdsourcing for survey research. Behav Res Methods 2011 Sep;43(3):800-813. [doi: 10.3758/s13428-011-0081-0] [Medline: 21437749]
58. Moder K. How to keep the Type I Error Rate in ANOVA if variances are heteroscedastic. Austr J Stat 2016 Apr 3;36(3):179. [doi: 10.17713/ajs.v36i3.329]
59. Games PA, Keselman HJ, Rogan JC. Simultaneous pairwise multiple comparison procedures for means when sample sizes are unequal. Psychol Bull 1981;90(3):594-598. [doi: 10.1037/0033-2909.90.3.594]
60. Liu Y, Tsyvinski A. National Bureau of Economic Research. 2018. Risks and Returns of Cryptocurrency URL: https://www.nber.org/papers/w24877 [accessed 2019-06-05] [WebCite Cache ID 78txGwdbY]
61. Tetsuya S. A microeconomic analysis of Bitcoin and illegal activities. In: Handbook of Digital Currency: Bitcoin, Innovation, Financial Instruments, and Big Data. Amsterdam, Netherlands: Elsevier Science; 2015:231-248.
62. Notheisen B, Hawlitschek F, Weinhardt C. Breaking down the blockchain hype–towards a blockchain market engineering approach. : Association for Information Systems; 2017 Presented at: Twenty-Fifth European Conference on Information Systems (ECIS); June 5-10, 2017; Guimarães, Portugal URL: https://pdfs.semanticscholar.org/9f21/9d266984bcf1589c77080439f06d49465c69.pdf
63. Connolly A, Kick A. What Differentiates Early Organization Adopters of Bitcoin From Non-Adopters? : AIS; 2015 Presented at: AMCIS 2015; August 13-15, 2015; Puerto Rico URL: https://aisel.aisnet.org/amcis2015/AdoptionofIT/GeneralPresentations/46/
64. Abramova S, Böhme R. Perceived benefit and risk as multidimensional determinants of bitcoin use: a quantitative exploratory study. : AIS; 2016 Presented at: ICIS 2016; December 11-14, 2016; Dublin, Ireland URL: https://aisel.aisnet.org/icis2016/Crowdsourcing/Presentations/19/
65. Möser M, Böhme R. Trends, tips, tolls: A longitudinal study of Bitcoin transaction fees. 2015 Presented at: International Conference on Financial Cryptography and Data Security; January 26-30, 2015; Berlin, Heidelberg p. 19-33.
66. Komiak S, Benbasat I. Understanding customer trust in agent-mediated electronic commerce, web-mediated electronic commerce, and traditional commerce. Inf Technol Manage 2004 Jan;5(1/2):181-207. [doi: 10.1023/B:ITEM.0000008081.55563.d4]
67. Yeager VA, Walker D, Cole E, Mora AM, Diana ML. Factors related to health information exchange participation and use. J Med Syst 2014 Aug;38(8):78. [doi: 10.1007/s10916-014-0078-1] [Medline: 24957395]
68. Venkatesh V, Morris M, Davis G, Davis F. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
69. Esmaeilzadeh P, Mirzaei T. Comparison of consumers' perspectives on different health information exchange (HIE) mechanisms: an experimental study. Int J Med Inform 2018 Dec;119:1-7. [doi: 10.1016/j.ijmedinf.2018.08.007] [Medline: 30342677]
70. Menachemi N, Rahurkar S, Harle CA, Vest JR. The benefits of health information exchange: an updated systematic review. J Am Med Inform Assoc 2018 Sep 1;25(9):1259-1265. [doi: 10.1093/jamia/ocy035] [Medline: 29718258]
71. Kuo TT, Ohno-Machado L. arXiv. 2018. ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks URL: https://arxiv.org/abs/1802.01746 [accessed 2019-06-03] [WebCite Cache ID 78rfyO3kp]
72. Linn, LA, Koo MB. ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland: HealthIT.gov; 2016. Blockchain for health data and its potential use in health it and health care related research URL: https://www.healthit.gov/sites/default/files/11-74-ablockchainforhealthcare.pdf [accessed 2019-06-05] [WebCite Cache ID 78tyZk5i3]
73. Unertl KM, Johnson KB, Lorenzi NM. Health information exchange technology on the front lines of healthcare: workflow factors and patterns of use. J Am Med Inform Assoc 2012;19(3):392-400 [FREE Full text] [doi: 10.1136/amiajnl-2011-000432] [Medline: 22003156]
74. Lam PE, Mitchell JC, Sundaram S. A formalization of HIPAA for a medical messaging system. 2009 Presented at: International Conference on Trust, Privacy and Security in Digital Business; September 3-4, 2009; Linz, Austria p. 73-85.
75. McDonald C. Protecting patients in health information exchange: a defense of the HIPAA privacy rule. Health Aff (Millwood) 2009;28(2):447-449 [FREE Full text] [doi: 10.1377/hlthaff.28.2.447] [Medline: 19276002]

XSL•FO
RenderX

76. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST. FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 2018;16:267-278 [FREE Full text] [doi: 10.1016/j.csbj.2018.07.004] [Medline: 30108685]
77. Prakash P. HealthIT.gov. 2016. Adoption of blockchain to enable the scalability and adoption of accountable care URL: https://www.healthit.gov/sites/default/files/13-71-blockchain_for_healthcare_paper_final.pdf [accessed 2019-06-03] [WebCite Cache ID 78rgl6q8J]
78. Ekblaw A, Azaria A, Halamka JD, Lippman A. A Case Study for Blockchain in Healthcare:"MedRec" prototype for electronic health records and medical research data. : IEEE; 2016 Presented at: 2nd International Conference on Open & Big Data 2016; August 22-24, 2016; Vienna, Austria URL: https://pdfs.semanticscholar.org/56e6/5b469cad2f3ebd560b3a10e7346780f4ab0a.pdf
79. Kiayias A, Panagiotakos G. Semantic Scholar. 2015. Speed-Security Tradeoffs in Blockchain Protocols URL: https://www.semanticscholar.org/paper/Speed-Security-Tradeo-s-in-Blockchain-Protocols-Kiayias-Panagiotakos/62646f9450a3c95e745c1d2bb056dcf851acdaad [accessed 2019-06-03] [WebCite Cache ID 78rh8DkFv]
80. Bradbury D. The problem with bitcoin. Comput Fraud Secur 2013 Nov;2013(11):5-8. [doi: 10.1016/S1361-3723(13)70101-5]
81. Ohno-Machado L, Bafna V, Boxwala AA, Chapman BE, Chapman WW, Chaudhuri K, iDASH team. iDASH: integrating data for analysis, anonymization, and sharing. J Am Med Inform Assoc 2012;19(2):196-201 [FREE Full text] [doi: 10.1136/amiajnl-2011-000538] [Medline: 22081224]
82. Dixon BE, Zafar A, Overhage JM. A framework for evaluating the costs, effort, and value of nationwide health information exchange. J Am Med Inform Assoc 2010;17(3):295-301 [FREE Full text] [doi: 10.1136/jamia.2009.000570] [Medline: 20442147]

Abbreviations

ANOVA: analysis of variance
e-commerce: electronic commerce
EHR: electronic health record
HIE: health information exchange
HIPAA: Health Insurance Portability and Accountability Act
IS: information systems
IT: information technology
TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology

Edited by P Zhang, K Clauson; submitted 28.03.19; peer-reviewed by C Esposito, M Kolotylo-Kulkarni; comments to author 04.05.19; revised version received 12.05.19; accepted 20.05.19; published 20.06.19.

Please cite as:
Esmaeilzadeh P, Mirzaei T
The Potential of Blockchain Technology for Health Information Exchange: Experimental Study From Patients' Perspectives
J Med Internet Res 2019;21(6):e14184
URL: http://www.jmir.org/2019/6/e14184/
doi:10.2196/14184
PMID:31223119

©Pouyan Esmaeilzadeh, Tala Mirzaei. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effects of Assistive Robot Behavior on Impressions of Patient Psychological Attributes: Vignette-Based Human-Robot Interaction Study

Meia Chita-Tegmark[1], PhD; Janet M Ackerman[1], BA; Matthias Scheutz[1], PhD

Tufts University, Medford, MA, United States

Corresponding Author:
Meia Chita-Tegmark, PhD
Tufts University
200 Boston Avenue #2530
Medford, MA, 02155
United States
Phone: 1 617 417 9090
Email: mihaela.chita_tegmark@tufts.edu

Abstract

Background: As robots are increasingly designed for health management applications, it is critical to not only consider the effects robots will have on patients but also consider a patient's wider social network, including the patient's caregivers and health care providers, among others.

Objective: In this paper we investigated how people evaluate robots that provide care and how they form impressions of the patient the robot cares for, based on how the robot represents the patient.

Methods: We have used a vignette-based study, showing participants hypothetical scenarios describing behaviors of assistive robots (patient-centered or task-centered) and measured their influence on people's evaluations of the robot itself (emotional intelligence [EI], trustworthiness, and acceptability) as well as people's perceptions of the patient for whom the robot provides care.

Results: We found that for scenarios describing a robot that acts in a patient-centered manner, the robot will not only be perceived as having higher EI (P=.003) but will also cause people to form more positive impressions of the patient that the robot cares for (P<.001). We replicated and expanded these results to other domains such as dieting, learning, and job training.

Conclusions: These results imply that robots could be used to enhance human-human relationships in the health care context and beyond.

(J Med Internet Res 2019;21(6):e13729) doi:10.2196/13729

KEYWORDS

robotics; emotional intelligence; patient-centered care

Introduction

With new advances in the fields of robotics and artificial intelligence, interest has grown in the introduction of robots as social agents in health care practice, especially for the management of chronic illness or geriatric conditions [1]. Socially assistive robots are machines designed to provide assistance through social means rather than physical ones, using social interactions for monitoring, coaching, providing companionship, and supporting health-promoting activities [2]. Robots are envisioned to play roles such as monitoring and record-keeping of symptom progression [3], helping with pill sorting and medication schedules [4], guiding people through therapeutic tasks [5], providing companionship [6], acting as stress reducers and mood enhancers [7], or supporting social interactions [8,9]. Although considerable attention has been given to the study of patient-robot interactions, less research has focused on the triangulation of patient-robot-others relationships, where *others* can be doctors, therapists, caregivers, or simply members of the society that the patient might interact with. As we design social robots for health care, we need to understand the effects these robots can have not only on the patient but also on the abovementioned *others* and how the robot fits overall into the network and dynamics of the patient's social relationships. This is important because social life and support

XSL•FO
RenderX

has been consistently shown to be a crucial predictor of health outcomes [10].

In this paper, we take the first steps toward studying the potential indirect effects of assistive social robots on the relationship between patients and others. We look at this through the lens of patient-centered care as a desirable approach [11] and emotional intelligence (EI) as a desirable set of capabilities for the robot [12-14].

In the first step of the study we investigated how robots can influence others' impressions of a patient's psychological attributes, asking, for example, whether people think of a patient as being competent, honest, and self-disciplined rather than disruptive, hostile, and disorganized. These psychological attributes have been shown to make a difference in the quality of care a patient may be given [15,16]. We proposed that people's impressions of a patient will be affected by the robot's behavior. Using text vignettes, we experimentally manipulated the robot's approach to care: *patient-centered* (focused on the needs and choices of the patient with regard to a treatment plan) or *task-centered* (focusing on how faithfully the treatment plan is being adhered to). We then investigated how the robot's approach influences: (a) people's perception of EI in the robot itself, (b) people's trust in the robot, (c) people's potential acceptance of the robot for the management of their own health, and finally (d) people's impressions of the patient. In step 2 of the study, we extended our investigation of social robots' influence on human-human relationships to other contexts: dieting, learning, and job training.

Human-Robot Interaction for Health Care Scenarios

Most studies of clinical applications of robots have focused on health outcomes for the patient (see [1]) or on robot acceptance by patients (see [17]). These studies are mainly concerned with the interaction between the patient and the robot, and only very few studies have investigated the effects the robot has on the patient's interactions with other people.

Several studies have looked at the social effects of PARO, a robot with the appearance of a baby seal that is responsive to touch, sound, temperature, and posture. By serving as the focus of the interaction between people (participants in these studies often interact in pairs or small groups with the robot), PARO was shown to have positive effects on social life such as increasing the density of social networks in a care home for the elderly [8], increasing social activity [18], and increasing social engagement of elderly nursing home residents with varying levels of dementia [19]. PARO was also shown to increase the reported quality of the interaction among people from a nonclinical population [20]. A couple of studies with robots used in therapy for children with autism have also been shown to have effects on the interaction between the child and others. In 1 study, the robot played a mediator role between the therapist and the child by allowing the child to express positive emotion in playing with the therapist [21]. In another study, the robot successfully served as a focus of free play between 2 children [22].

In nonclinical contexts, the social mediation effect of robots has been studied, among others, for the following purposes: conflict resolution with children [23] or adults [24,25], active listening [26], teaching EI through interactive storytelling [27], and enhancing cooperation [28]. Although these results represent a promising start, more research is needed to better understand both the ways in which robots can support social interactions among people and how they might inadvertently influence human-human interactions in perhaps undesirable ways. In clinical contexts, different aspects of human-human interactions, for example, how patients are perceived by others, can have important implications for health care.

How the Impression of Patient Psychological Attributes Affects Care Decisions

A number of studies have documented the effects of a patient's affect or perceived personality traits on care decisions; in short, doctors appear willing to prescribe more care for more positive or likable patients. Despite only ranking emotional state as an important consideration in decisions about Intensive Care Unit admission 6% of the time, doctors were almost 3 times more likely to recommend admission of a hypothetical patient if given a vignette that described the patient as upbeat and courageous rather than sad and discouraged [15]. Similarly, another set of doctors recommended more follow-up visits and calls for simulated likable and competent patients and more staff time spent educating simulated likable patients [16].

Other research has indicated that doctors have more positive feelings toward patients who appear happy [29,30] or toward those who express both positive and negative affect over the course of a visit [31]. On the contrary, some primary care physicians report that their challenging patients often become favorites over time and that favorite patients likely benefit from the extra effort that physicians feel inclined to spend on these patients' care [32].

Doctors are not the only ones who are influenced in their behavior toward patients by the patient's perceived attributes. A study found that participants (who were not specifically medical professionals) tend to dislike patients who appear distraught but are slightly more willing to aid patients displaying negative affect than those displaying positive affect, offering the least aid to those who show little affect [33].

Given how consequential people's perceptions of patient psychological attributes are to the patients receiving optimal care, roboticists should be mindful of how they design robots so as not to negatively affect relationships between patients and health care providers or other people. This could also be seen as an opportunity for health care robotics: Social robots could be used to enhance the perception of positive attributes and promote good relationships among patients and doctors, caregivers, or others.

Patient-Centered Care and Emotional Intelligence as Guidelines for Robot Design

How robots influence relationships in the context of health care, whether in a desirable or undesirable way, will of course depend on their design. More specifically, it will depend on what approach the robot will take for providing care to the patient, and what social capabilities the robot will have.

XSL•FO
RenderX

Patient-centered care [34] is an influential approach to health care that has formulated desirable features for the relationship of the patient with others. This approach emphasizes respect for the patient's preferences, information, education and communication with the patient, coordination of care, emotional support for the patient, physical comfort, involvement of family, continuity and transition, and access to care [35]. It has been proposed that success in providing patient-centered care may depend on social capabilities such as EI [36].

EI has been linked to positive social relationships in multiple contexts, including health care. EI comprises abilities such as perceiving, understanding, and managing one's own emotions and the emotions of others. EI has been linked to benefits for both the person possessing these abilities, such as enhanced job performance and stress management [13,37] and better educational outcomes [38], and also for the social group that one is embedded in: EI has been associated with improved teamwork and conflict resolution [39], higher leadership ratings [40], and successful social interactions [41].

In the health care context, EI in medicine has been linked to positive doctor-patient relationships, increased empathy, better teamwork and communication skills, stress management, and organizational commitment and leadership [14]. EI concepts are also central to nursing practice, with implications for nursing students' learning, ethical decision making, critical thinking, evidence, and knowledge use in practice (for a review see [42]). We thus proposed that EI is needed in social robots that are to operate in the health care setting. To our knowledge, only a couple of studies have so far investigated EI in robots [43,44] and have found that people do expect and detect differences in EI in robot agents and that these differences influence how much robots are trusted [43]. It is thus important to understand people's perceptions of robots' EI in the context of patient care.

Social Robots and Human-Human Relationships Beyond the Health Care Context

Patient care, however, is not the only context in which robots' influence on relationships between people can have an impact on health and well-being. In the domain of public health, robots could be meaningfully used to mitigate problems such as epidemic obesity and to enhance the relationship between clients and dieticians or weight loss support groups. In the field of education, robots could help promote individualized learning plans and student-centered approaches as well as enhance relationships between students and teachers. In the industry, robots could optimize mentor-trainee interactions and assist in retraining people for new jobs, which has been increasingly needed as technological advances are reshaping the labor market. We have focused the first step of this study on how robots affect relationships in the health care setting, and then we broadened the scope of our inquiry in the second step of the study to other contexts that are more indirectly linked to people's health and well-being.

Aims of the Current Study

The aims of this study were to investigate the effects of robot behavior in the health care context and to probe whether similar effects extend to other contexts relevant for well-being and quality of life. In the first step of the study, we investigated whether the robot behavior, patient-centered or task-centered, has an impact on (a) how the robot is perceived, in terms of EI, trustworthiness, and acceptability and (b) the impression that people form of the patient assisted by the robot. We hypothesized that (1) a robot that acts in a patient-centered way will be perceived as having higher EI, (2) inspired by the findings of [43], a robot that acts in a patient-centered way will be trusted more, and (3) accepted more, and we also hypothesized that (4) by acting in a patient-centered way and respecting the patient's agency, the robot will cause others to think more highly of the patient.

In the second step of the study we further investigated the extent to which the effects observed in the health care context can be replicated and extended to other contexts relevant for people's health and well-being, contexts in which social robots can provide assistance by monitoring, keeping record, and informing a professional about the user's progress.

Methods

Design

The experiment used a between-group design and a text vignette methodology (see the Materials section). For the first step of the study, investigating the effects of robot behavior in the health care context, we conducted a series of one-way analyses of variance (ANOVA). Our dependent variables were perceptions of robot EI, trust in the robot, robot acceptance, and impression of the patient cared for by the robot (see the Measures section). Our main independent variable was the condition based on which the vignette described the robot's behavior as being either patient-centered or task-centered (see the Materials section for further details). We also verified for effects of gender as an independent variable and age as a covariate in a series of analyses covariance (ANCOVA).

For the second step of the study, investigating the effects of robot behavior in other contexts relevant for well-being, we used a 2 × 3 between-group experimental design. In addition to the 2 conditions (person-centered or task-centered robot behavior), we also designed vignettes that varied in context (weight loss or learning or job training), which we used as an additional independent variable. As for the health care context, scores from our 4 different questionnaires were used as dependent variables. Gender and participant age were added as variables in the models in a further step.

Our participants were recruited on the Amazon Mechanical Turk (AMT) platform. This made it possible to reach participants from more diverse ethnic backgrounds and spanning a wider age range than what is typical for in-laboratory studies. To ensure language comprehension and a similar compensation incentive, we only recruited participants based in the United States who were fluent in English.

Participants

For the first step of the study, in which we investigated the effect of robot behavior in the health care context, a total of 199 participants completed the experiment through the AMT. A total of 11 participants failed to pass our attention checks and

XSL•FO
RenderX

were excluded from the analyses. According to standard practice for AMT studies, our attention check consisted of a reading comprehension question about the topic of the vignette meant to assess whether participants read it attentively. The 188 participants with usable data ranged in age from 20 to 72 (mean 34.58, SD 10.04) years, 77 were female, and 2 identified as Other. The ethnic composition of the sample was white or Caucasian (74.5%, n=140), Asian (5.8%, n=11), African American (9%, n=17), Hispanic (6.9%, n=13), and Other (3.7%, n=7). In total, 95 participants saw the person-centered vignette and 93 participants saw the task-centered one.

The second step of the study, in which we investigated the effects of robot behavior in other contexts related to well-being, was completed by 299 participants, out of which 254 passed our attention checks, and thus provided usable data. Participants' ages ranged from 20 to 71 (mean 36.48, SD 10.54) years, and 111 of the participants were female and 5 of the participants identified as *Other*. The ethnic composition of the sample was as follows: white or Caucasian (79.1%, n=201), Asian (3.1%, n=8), African American (7.8%, n=20), Hispanic (6.7%, n=17), and Other (3.1%, n=8). A total of 93 participants saw the weight loss vignette, 92 saw the learning vignette, and 69 saw the job training vignette. The procedure was identical to that in step 1 but with more vignettes. Each participant saw one vignette only, either in the person-centered or task-centered form.

Measures

Perceptions of Robot Emotional Intelligence

We used a 24-item questionnaire based on a measure developed by Caruso and Salovey [45] and previously used by Fan et al. [43] to measure perceptions of EI in robots. Items referred to emotion perception (eg, "Knows why people feel the way they do"), understanding (eg, "Considerate of others' feelings"), and management (eg, "Creates positive moods in people") capacities and participants indicated, on a 5-point Likert scale, how much each statement described the robot in the vignette from *0=not at all* to *4=very much so* (Cronbach alpha=.96). We averaged the scores from all the items for each participant.

Trust in the Robot

We adapted a 4-item, 5-point Likert scale measure from a study by Mayer and Davis [46] to match the context of the vignette. An example item is "I would be willing to let this robot have control over my health care management." The measure was adapted for the second step of the study, for example items would refer to weight loss plan instead of health care management for the dieting vignette.

Robot Acceptance

Participants rated 2 statements on a 5-point Likert scale from *0=strongly disagree* to *4=strongly agree* about finding the robot useful ("If I were chronically ill, I would find it useful to have a robot like this to help with my treatment") and wanting to use a similar robot to the one in the vignette ("If I were chronically ill, I would want to use a robot like this to help with my treatment"), should they find themselves in a similar situation to that of the patient. The measure was adapted to match the other contexts for the second step of the study.

Impression of Patient

We developed a measure based on the literature investigating the impressions that health care providers have of patient psychological attributes and how that affects their care decisions [15,16,29,47,48]. A set of relevant patient descriptors were collected: dependable and self-disciplined, disorganized and careless, capable of participating in treatment and adhering to health care recommendations, likable, competent, having a positive attitude, defiant, disruptive, hostile, and honest. Participants rated 10 statements about the patient formulated around the descriptors above (eg, "I feel the patient is disorganized and careless") on a 5-point Likert scale from "0 = strongly disagree" to "4 = strongly agree" (Cronbach alpha=.81). Items indicating negative psychological attributes were reverse-scored. We averaged the scores from all the items for each participant. For consistency and ease of comparison, the person descriptors used in this measure were not changed for step 2 of the study. Note that even though these items were constructed based on the literature on impressions of patient psychological attributes, they are relevant across contexts.

Materials

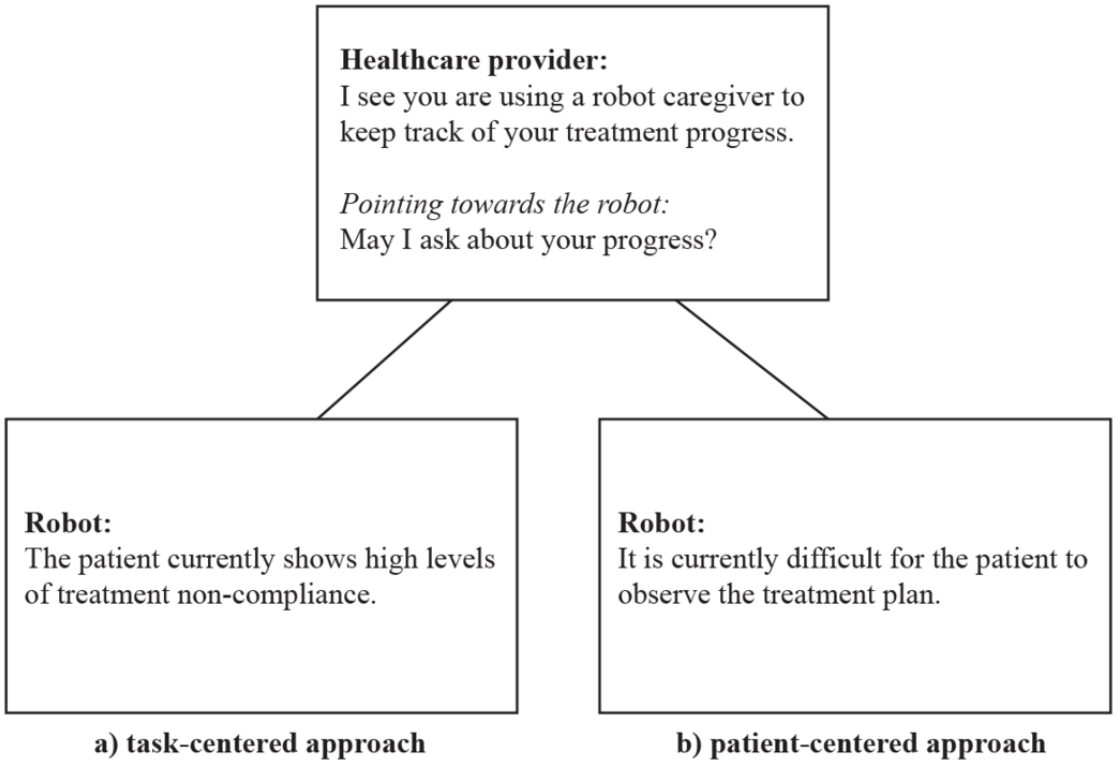
For this study, we used the text vignette methodology to evaluate people's perceptions, attitudes, and impressions toward assistive robots, their behavior, and the patients they care for. In text vignettes, hypothetical situations are described to which participants are asked to respond. This is a common methodology in psychology and sociology experiments and has been used successfully in human-robot interaction research (eg, [43]) as an initial step of investigating and informing possible robot design choices.

For the first step of the study that focuses on robot behavior in the health care context we used a text vignette featuring 3 characters: a health care provider, a patient, and an assistive robot. The scenario starts with the health care provider noticing that the patient is using an assistive robot for treatment monitoring and asks permission to get a report from the robot about treatment progress, to which the patient responds affirmatively. In one condition, the robot gives a patient-centered report that focuses on the needs and choices of the patient with regard to the treatment plan (patient-centered condition). In the other condition, the robot gives a task-centered report focused on how faithfully the treatment plan was adhered to (task-centered condition; see Figure 1). The treatment plan refers to a medication schedule and a physical exercise routine.

The vignette was presented in a video format. The dialogue was parsed into chunks usually containing one dialogue turn or, in the case of the robot, which speaks for longer into sentences that conveyed one idea. Each chunk was displayed on a blank screen. The text appeared across the screen as if it was being typed (using the typewriter animation effect) with a speed similar to what it would take to say the words out loud. This gave the impression of a dynamic conversation. The speaker was clearly indicated at the beginning of each dialogue turn or chunk of text (eg, "Robot: The patient...").

XSL•FO
RenderX

Figure 1. Sample screenshots illustrating the vignette with the 2 conditions: (a) task-centered and (b) patient-centered.



The vignette characters were referred to as *health care provider*, *patient*, and *robot*, and no personal characteristics were specified about the human characters (eg, no gender was suggested). In addition, no description of the robot's capabilities or appearance (eg, whether it was humanoid or not) was provided; rather, this was left to the participants' imagination. From the vignette, it could be inferred that the robot had the capacity to track treatment progress and produce a spoken report.

For the second step of the study in which we sought to generalize our findings to other contexts, 3 additional vignettes (with 2 conditions each) were used, modeled closely after the vignette in step 1 in terms of (a) structure (ie, the same number of characters and the same conversation progression were kept), (b) theme (ie, the robot monitored progress with regard to a body-oriented task such as dieting and exercising, dancing, and training for a physically active job), (c) conditions (ie, an assistant robot gave a report in a person-centered or task-centered manner), (d) display (ie, the same procedures for chunking and displaying text were used), and (e) amount of information conveyed (ie, no personal characteristics of the humans and no description of the robot were given). The vignettes differed in the context described and the characters involved.

In the first vignette (the dieting vignette), the robot assistant was used for monitoring weight loss progress and the interaction happened between a client, a weight loss coach, and the assistant robot. The weight loss plan involved a meal plan and a physical exercise routine. The second vignette (the learning vignette) featured a robot assistant that was used to keep track of progress in learning how to dance. The interaction took place between a student, a dance instructor, and the assistant robot. The learning goals involved a dance practice schedule and a strength and conditioning routine. Finally, in the third vignette (training vignette), a robot assistant was used for monitoring training progress for volunteer firefighters. The interaction involved a volunteer firefighter, a lieutenant, and the assistant robot. The training goals involved a physical exercise training plan and *Search and Rescue* drill training. Given the diversity of contexts and characters in this second step, we have renamed the *patient-centered condition* from step 1 as the *person-centered condition*, *person* here referring to the weight loss client, the student, or the volunteer firefighter.

Procedure

Participants accessed the experiment through the AMT. After reading and agreeing to the consent form, participants filled out a demographic questionnaire (age, gender, and ethnicity). They then followed instructions for watching the video. Participants were randomly assigned to one condition: person-centered or task-centered. After watching the video, participants completed the measures in the following order: perception of robot's EI, trust in the robot, robot acceptance, and impressions of the patient. After concluding the experiment, participants were able to collect their US $1 compensation. All procedures were reviewed and approved by the Institutional Review Board at our university.

Results

Perceptions of Robot Emotional Intelligence

Step 1: Robot Behavior in Health Care

We began by investigating whether participants perceived a difference in the robot's EI depending on whether the robot gave a patient-centered or a task-centered report (see the left side of Figure 2 and Table 1).

XSL•FO
RenderX

Figure 2. Effect of conditions on (a) perceptions of robot's emotional intelligence (EI) and (b) impressions of patient (mean and SE).

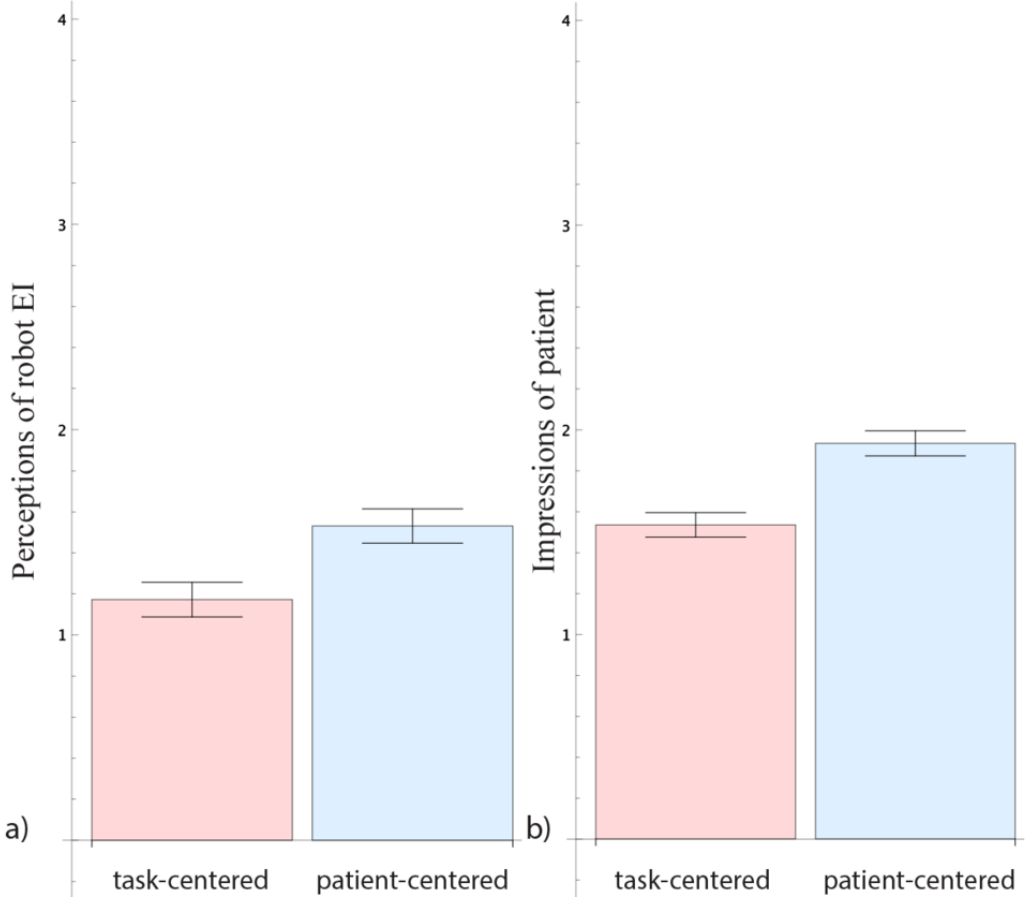


Table 1. Descriptive statistics and ANOVA summaries for the *condition* variable.

Dependent variables	Condition				Statistics		
	Task-centric		Patient-centric				
	N	Mean (SE)	N	Mean (SE)	*F* (df)	*P* value	
Perceptions of robot emotional intelligence	92	1.17 (0.08)	93	1.53 (0.08)	9.14 (1)	.003	0.05
Trust in the robot	93	1.8 (0.09)	95	1.97 (0.08)	1.84 (1)	.18	0.01
Robot acceptance	92	2.15 (0.13)	95	2.33 (0.12)	1.01 (1)	.32	<0.01
Impression of patient	91	1.53 (0.06)	92	1.93 (0.06)	21.62 (1)	<.001	0.11

We conducted a one-way ANOVA with the *perception of robot EI* as a dependent variable and *condition (patient-centered/task-centered)* as an independent variable. We found a significant effect of *condition*, with the EI of the robot rated significantly higher in the patient-centered condition ($F_{1,183}$= 9.14; *P*=.003; and =0.05). We also conducted an ANCOVA, adding the participant's *gender* (male/female) as an independent variable besides *condition* and the participant's *age* as a covariate. We again found a significant effect of *condition* ($F_{1,178}$=5.99; *P*<.001; and =0.05). No significant effect of participant *gender* or significant interaction between *condition* and *gender* was found. Participant's *age*, however, had a significant influence on *perceptions of robot EI*, with older participants rating the robot as having lower EI ($F_{1,178}$=12.89; *P*<.001; and =0.07). This is consistent with the well-known effect of older adults having less favorable opinions of robots in general.

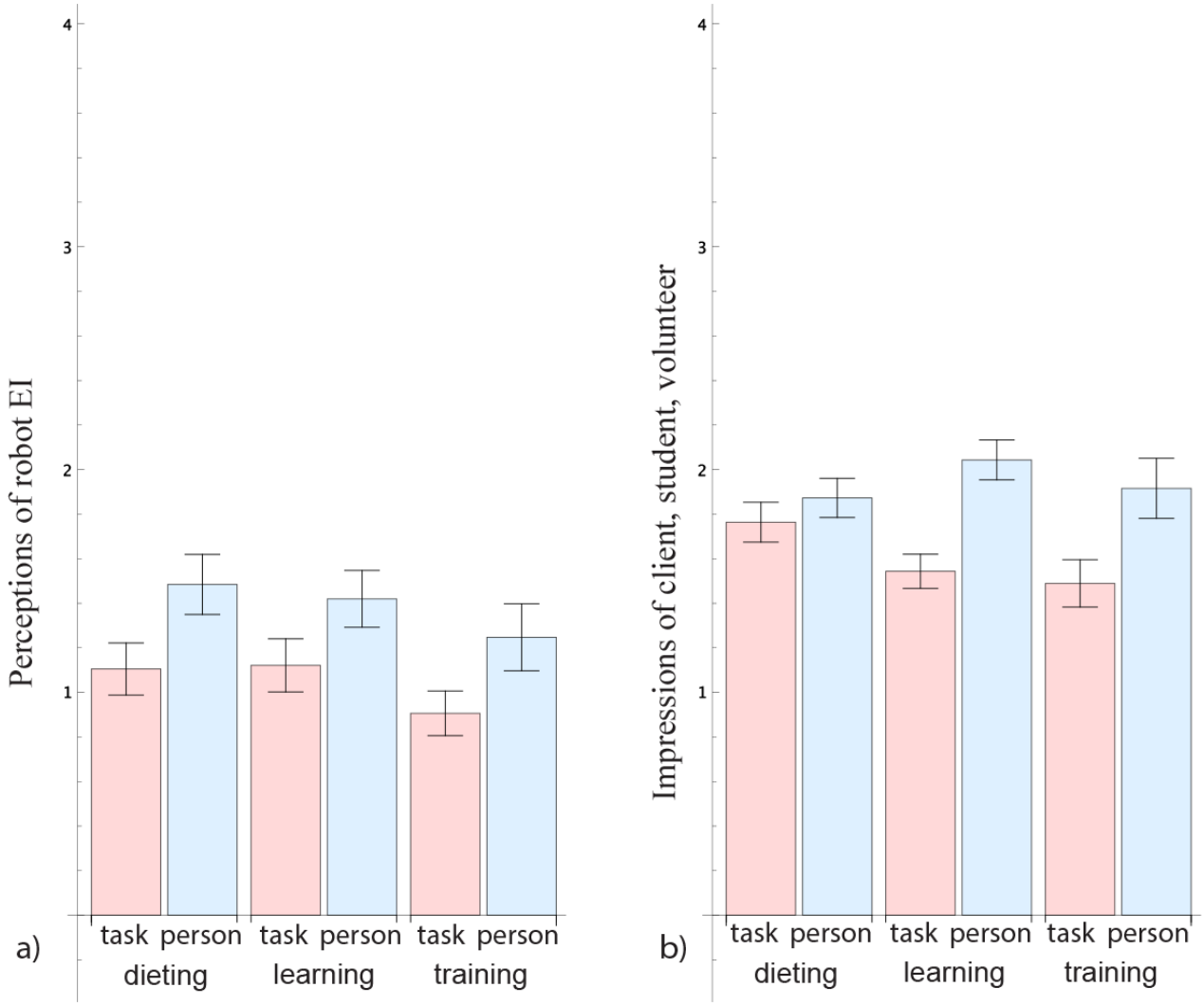
Step 2: Robot Behavior in Other Contexts

Mirroring the first step, we began by investigating whether participants perceived a difference in the robot's EI based on the condition they were exposed to and the type of context described by the vignette (see left side of Figure 3 and Table 2). We conducted a 2 × 3 ANOVA with *perceptions of robot EI* as the dependent variable and *context (weight loss/learning/job training)* and *condition (person-centered/task-centered)* as independent variables (see Table 3). We found a significant main effect of *condition*, with participants perceiving higher EI when the robot gave a person-centered report $F_{1,243}$=10.71, *P*=.001, and =0.04. The *context* had no significant effect on *perceptions of robot EI*. There was also no

XSL•FO
RenderX

significant interaction between the *context* and the *condition*. We also conducted an ANCOVA adding participant *gender* as an independent variable besides *condition* and *context* and participants' *age* as a covariate. Similar to step 1, we found a significant main effect of *condition* ($F_{1,231}$=7.56, P<.001, and [illegible]=0.04) and *age* ($F_{1,231}$=5.53, P=.02, and [illegible]=0.02) but no other significant main effects or interactions.

Figure 3. Effects of conditions (task-centered vs person-centered) on (a) perceptions of robot emotional intelligence and (b) impressions of client, student, and volunteer in dieting, learning, and training contexts (mean and SE). Note that for the dieting context, the error bars corresponding to the mean ratings of impression of the client are overlapping between the 2 conditions. Although the same trend is seen as in the other contexts, with higher ratings (ie, more positive impressions) in the person-centered condition, in the case of dieting taken separately, the difference between the conditions is not significant. We list potential reasons in the Discussion section.



XSL•FO
RenderX

Table 2. Descriptive statistics for the variables *condition* and *context*.

Condition and context	N	Mean (SE)
Perceptions of robot EI		
Task-centric		
Weight loss	48	1.1 (0.12)
Learning	47	1.12 (0.12)
Job training	36	0.91 (0.10)
Person-centric		
Weight loss	43	1.48 (0.13)
Learning	44	1.42 (0.13)
Job training	31	1.25 (0.15)
Trust in the robot		
Task-centric		
Weight loss	48	1.97 (0.14)
Learning	47	1.93 (0.11)
Job training	37	1.62 (0.18)
Person-centric		
Weight loss	43	2.17 (0.13)
Learning	44	2.04 (0.13)
Job training	32	2.02 (0.17)
Robot acceptance		
Task-centric		
Weight loss	48	1.87 (0.21)
Learning	47	2.31 (0.17)
Job training	37	1.69 (0.23)
Person-centric		
Weight loss	44	2.47 (0.17)
Learning	45	2.23 (0.19)
Job training	31	2.27 (0.22)
Impression of patient		
Task-centric		
Weight loss	47	1.76 (0.09)
Learning	46	1.54 (0.08)
Job training	37	1.49 (0.11)
Person-centric		
Weight loss	44	1.87 (0.09)
Learning	44	2.04 (0.09)
Job training	32	1.92 (0.13)

XSL•FO
RenderX

Table 3. ANOVA summaries for the variables *condition* and *context*.

Variables	Mean squares	*F* (df)	*P* value	
Perceptions of robot emotional intelligence				
Condition	7.00	10.71 (1)	.001	0.04
Context	1.05	1.60 (2)	.2	0.01
Condition × context	0.04	0.06 (2)	.94	<0.01
Residual	0.65	—[a] (243)	—	—
Total	0.68	— (248)	—	—
Trust in the robot				
Condition	3.51	4.14 (1)	.04	0.02
Context	1.22	1.45 (2)	.24	0.01
Condition × context	0.42	0.50 (2)	.61	<0.01
Residual	0.85	— (245)	—	—
Total	0.86	— (250)	—	—
Robot acceptance				
Condition	8.28	5.00 (1)	.03	0.02
Context	1.64	0.99 (2)	.37	<0.01
Condition × context	3.21	1.94 (2)	.15	0.01
Residual	1.65	— (246)	—	—
Total	1.69	— (251)	—	—
Impression of person				
Condition	7.30	19.30 (1)	<.001	0.07
Context	0.28	0.74 (2)	.48	<0.01
Condition × context	0.96	2.53 (2)	.08	0.02
Residual	0.38	— (244)	—	—
Total	0.41	— (249)	—	—

[a]Not applicable.

Trust in the Robot

Step 1: Robot Behavior in Health Care

To investigate whether the robot's patient-centered or task-centered approach had an influence on how much people trusted the robot, we conducted a one-way ANOVA with *trust in the robot* as the dependent variable and *condition* (*patient-centered/task-centered*) as an independent variable. We found no significant effects of *condition* on *trust in the robot*. We further conducted an ANCOVA with participants' *gender* as an additional independent variable and the participants' *age* as a covariate, finding no significant effects of *gender* or *age* on *trust in the robot*.

Step 2: Robot Behavior in Other Contexts

We then investigated whether trust in the robot was affected by the person-centered or task-centered conditions or by the vignette context (weight loss/learning/job training). We did this by conducting a 2 × 3 ANOVA with *trust in the robot* as the dependent variable and *condition* and *context* as independent variables. The main effect of *condition (person-centered/task-centered)* was significant $F_{1,245}$=4.14, *P*=.04, and =0.02, with participants in the person-centered condition showing more trust in the robot. An ANCOVA with participants' *gender* and participants' *age* added as variables showed no additional main effects. For the ANCOVA, the main effect of condition did not meet 5% significance levels $F_{1,233}$=3.52, *P*=.06, and =0.01.

Robot Acceptance

Step 1: Robot Behavior in Health Care

Next, we explored the potential effect of the robot's care approach on how helpful and desirable participants would find using such a robot for their own health management. We conducted a one-way ANOVA with *robot acceptance* as the dependent variable and *condition* (*patient-centered/task-centered*) as an independent variable. We found no significant effects of *condition* on *robot acceptance*. An ANCOVA by adding participant *gender* as an independent variable and participant *age* as a covariate yielded no further significant findings.

XSL•FO
RenderX

Step 2: Robot Behavior in Other Contexts

To explore the influence of the 2 conditions as well as that of the vignette context (weight loss/learning/job training) on robot acceptance (participants finding the robot potentially helpful and desirable to use), we conducted another 2 × 3 ANOVA, this time with the *robot acceptance* measure as the dependent variable and *condition* and *context* as independent variables. We again found a main effect of *condition*, with *robot acceptance* being significantly higher for the condition in which the robot provided a person-centered report, $F_{1,246}$=5.00, P=.03, and =0.02. As for the other measures, we also conducted an ANCOVA with participants' *gender* and *age* as additional variables and found no significant main or interaction effects other than that of *condition*$F_{1,234}$=4.67, P=.03, and =0.01.

Impression of Patient

Step 1: Robot Behavior in Health Care

We also examined the effect of the robot's patient-centered or task-centered behavior on participants' impressions of the patient's psychological attributes (see right side of Figure 2) by conducting a one-way ANOVA with the *impression of patient* as a dependent variable and *condition* as an independent variable. We found a significant effect of *condition* on the *impression of patient*, with participants exposed to the patient-centered robot reporting a higher *impression of the patient* score (ie, having a more favorable impression; $F_{1,181}$=21.62; P<.001; and =0.11). We also conducted an ANCOVA with participant *gender* and the interaction between *gender* and *condition* as additional independent variables and participant *age* as a covariate. We again found a significant main effect of *condition*, $F_{1,176}$=19.13; P<.001; and =0.10, and further a significant main effect of *gender*, $F_{1,176}$=4.01, P=.05, =0.02, with female participants rating their impressions of the patient more positively.

Step 2: Robot Behavior in Other Contexts

Finally, we examined the influence of the 2 conditions (person-centered and task-centered robot report) and the context of the vignettes (dieting/learning/job training) on the participants' impressions of the person (the client, the student, and the volunteer; see Figure 3). We conducted a 2 × 3 ANOVA with the *impression of the person* as the dependent variable and *condition* and *context* as the independent variables. We found a main effect of *condition*$F_{1,244}$=19.30, P<.001, and =0.07. Participants in the person-centered condition rated their impression of the client, student, or volunteer as significantly higher (more positive). As before, we proceeded to conducting an ANCOVA, adding participants' *gender* as an independent variable and participants' *age* as a covariate. We found no significant main or interaction effects of *gender* or *age* but *condition* emerged again as significant $F_{1,246}$=16.16, P<.001, and =0.06.

Discussion

For the health care context, our findings confirm our first hypothesis, namely that robots that exhibit patient-centered behavior are perceived as having higher EI. However, we found no support for our second and third hypotheses: there was no indication that people trusted or accepted the robot less when it behaved in a task-centered way. When Fan et al. [43] found differences in trust based on the robots' level of EI, their experimental manipulation was much more extreme than the one in this study: They compared a robot that was polite and understanding with a robot that was rude and abrasive.

Also, our results show that the robot's behavior was able to significantly influence people's impressions of the patient. It is striking that the robot modeling respect for the patient's agency positively influenced how people viewed the patient. This is particularly remarkable as the facts reported by the robot were identical: in both conditions, the patient is reported to have failed the same number of times in respecting the medication schedule and following the exercise routine. This suggests new opportunities for using robots to positively influence care for patients.

Beyond health care, we have replicated these results in other contexts relevant to well-being, which shows that these findings generalize to other circumstances. This suggests that social robots' influence on human-human interactions are not specific to medical treatment adherence but generalize to other categories of interactions. Our findings from step 2 of the study show that robots are perceived as being more emotionally intelligent when they behave in a person-centric way and that their behavior can influence how other people perceive the person they are assisting. In the second step of the study, but not the first, we also found small effects of the robot's behavior on how trusted or accepted the robot was.

To our knowledge, this is the first investigation that probed the potential influences of assistive robots on how patients are judged by others. By simply varying the way the robot gives a treatment progress report (manipulating the language used, but not the facts conveyed), the robot makes a significant difference to the extent to which the patient is thought of as likable, self-disciplined, competent, and having a positive attitude as opposed to being hostile, disruptive, and disorganized. People take cues from the robot, and when the robot uses a patient-centered language that indicates respect for the patient's agency, as opposed to a task-centered language that focuses on the treatment benchmarks, people form a more positive impression of the patient.

It is remarkable that robots are able to have this effect given that their language output is scripted and certainly does not come with the emotional connotations that a person's choice of language would have. A robot's behavior is not connected to beliefs and attitudes in the same way a human's behavior is, yet our findings suggest that we inadvertently let our perceptions and impressions of others be guided by the robot's actions. This can be seen both as having cautionary implications for the field of human-robot interaction but also as an opportunity for the field: if robots can have an influence on how we think of others,

XSL•FO
RenderX

then perhaps they can be used for improving relationships in the health care setting and beyond.

Although these results suggest promising possibilities, much more research is needed to understand how assistive social robots can be optimally embedded in social interactions in the health care context. For starters, several gaps and limitations of this study need to be addressed by further research. Because the study is based on hypothetical vignettes, caution is needed in generalizing these findings to real-world interactions. Also, participants in our experiments were laypeople with no connection to the patient. It is not clear whether health care providers or people invested in the patient's health will be equally susceptible to influence from robots. In addition, participants themselves might have had different amounts of experience with giving care and support to sick individuals, and that might influence how susceptible they are to being influenced by the robot. Also, their own health status might have a bearing on that. Someone struggling with their own health and treatment might be more sympathetic to another person's need for agency over their own treatment and that might modulate the robot's influence. It is important to note that variations in cultural norms regarding the care of the sick people, beliefs about how much privacy, agency, and responsibility the patient should have versus the caregivers, will likely influence how people perceive the robot's behavior as well. The potential interaction of these factors with the effects of the robot behavior remains to be determined through further research.

Another limitation of our study is that the impressions of the patient were formed in the absence of any other information about the patient. In real-life scenarios, even for very short, first-time interactions, people give off an abundance of signals about who they are (age, gender, ethnic group, and social class) and research has shown that humans are able to form impressions extraordinarily fast [49]. It is therefore likely that the robot's effect on people's impression might be greatly modulated by the information conveyed by the patient herself. For example, people might hold different beliefs to start with in terms of how much agency someone should have over their own treatment based on whether they are an older adult versus a young adult versus a child. In our vignettes, people's characteristics were left completely to the participant's imagination, and some might have formed very vivid images of the patient whereas others less so.

Similarly, the robot characteristics were left to the imagination of the participants, which might have contributed to how the robot was perceived. The contributions of various characteristics to the robot perception are likely mixed. For example, the findings of Fan et al [43] suggest that some characteristics are very resistant to changes in the perception of EI in robots: the robot's voice and gestures did not have an influence on people's ratings of the robot's EI, the authors finding no differences between video, audio, or text vignettes. It seems that what the robot says is the single most important factor in how emotionally intelligent the robot is perceived. However, another study [44] found that the robot assigned gender (indicated through name), in addition to what the robot said influenced how emotionally intelligent the robot was perceived to be.

Finally, although we found that the robot's influence on others' impression of the person the robot is assisting is generalizable to a variety of contexts, the magnitude of the effect might vary from context to context. Figure 3 suggests that people are not equally susceptible to having their impressions influenced by the robot's behavior in the dieting context. For dieting, the task-centered behavior of the robot does not negatively affect people's impression of the person trying to lose weight. Perhaps people have more experience with dieting and how difficult it can be to adhere to a dieting plan, and thus feel more empathy toward the weight loss coaching client. Another possibility is that people already feel that individuals should have agency and choices over their weight, much more so than over their medical treatment, studies, or job performance. Further research is needed to determine the context-related factors that influence how people form their opinions of others.

Conclusions

This is the first study, to our knowledge, to investigate the influence of robots' behavior on the impressions that people form about the person the robot is assisting. We found that people perceived the robot as having higher EI when it was behaving in a person-centric way and in some contexts as being more trustworthy and having a higher acceptance rate. Most importantly, robots were able to influence people's impressions of patients, coaching clients, students, and volunteers by modeling behavior that was respectful of these people's agency and choices with regard to medical treatment, weight loss plans, learning, and job training. The most immediate implications for human-robot interaction research are that (1) it is important to be aware of the social assistive robots' influence on the relationships between the people assisted and others and (2) the ability to positively impact human relationships opens up new exciting opportunities for the use of robots in pursuing health and well-being.

Acknowledgments

This project was supported by the National Science Foundation grant IIS-1316809.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Abdi J, Al-Hindawi A, Ng T, Vizcaychipi MP. Scoping review on the use of socially assistive robot technology in elderly care. BMJ Open 2018 Dec 12;8(2):e018815 [FREE Full text] [doi: 10.1136/bmjopen-2017-018815] [Medline: 29440212]
2. Matarić MJ. Socially assistive robotics: human augmentation versus automation. Sci Robot 2017 Mar 15;2(4):eaam5410. [doi: 10.1126/scirobotics.aam5410]
3. Briggs P, Scheutz M, Tickle-Degnen L. Are robots ready for administering health status surveys':First results from an HRI study with Parkinson's Disease. In: Proceedings of the Tenth Annual ACM/IEEE International Conference for Human-Robot Interaction. 2015 Mar Presented at: HRI'15; March 2-5, 2015; Portland, USA p. 327-334. [doi: 10.1145/2696454.2696476]
4. Wilson JR, Tickle-Degnen L, Scheutz M. Designing a social robot to assist in medication sorting. In: Agah A, Cabibihan JJ, Howard A, Salichs M, He H, editors. Social Robotics. Cham: Springer; 2016:211-221.
5. Kim GH, Jeon S, Im K, Kwon H, Lee BH, Kim GY, et al. Structural brain changes after traditional and robot-assisted multi-domain cognitive training in community-dwelling healthy elderly. PLoS One 2015;10(4):e0123251 [FREE Full text] [doi: 10.1371/journal.pone.0123251] [Medline: 25898367]
6. Banks MR, Willoughby LM, Banks WA. Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J Am Med Dir Assoc 2008 Mar;9(3):173-177. [doi: 10.1016/j.jamda.2007.11.007] [Medline: 18294600]
7. Takayanagi K, Kirita T, Shibata T. Comparison of verbal and emotional responses of elderly people with mild/moderate dementia and those with severe dementia in responses to seal robot, PARO. Front Aging Neurosci 2014;6:257 [FREE Full text] [doi: 10.3389/fnagi.2014.00257] [Medline: 25309434]
8. Wada K, Takanori S. Social effects of robot therapy in a care-house - change of social network of the residents for two months. In: Proceedings 2007 IEEE International Conference on Robotics and Automation. 2007 Presented at: 2007 IEEE International Conference on Robotics and Automation; April 10-14, 2007; Rome, Italy p. 1250-1255. [doi: 10.1109/ROBOT.2007.363156]
9. Arkin RC, Scheutz M, Tickle-Degnen L. Preserving dignity in patient caregiver relationships using moral emotions and robots. 2014 Presented at: 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering; May 13-14, 2014; Chicago, USA p. 1-5.
10. Uchino BN. Social support and health: a review of physiological processes potentially underlying links to disease outcomes. J Behav Med 2006 Aug;29(4):377-387. [doi: 10.1007/s10865-006-9056-5] [Medline: 16758315]
11. Stewart M. Towards a global definition of patient centred care. Br Med J 2001 Feb 24;322(7284):444-445 [FREE Full text] [doi: 10.1136/bmj.322.7284.444] [Medline: 11222407]
12. Salovey P, Mayer J. Emotional intelligence. Imagin Cogn Pers 2016 Nov 28;9(3):185-211. [doi: 10.2190/DUGG-P24E-52WK-6CDG]
13. Goleman D. Emotional Intelligence: Why It Can Matter More Than IQ. London, United Kingdom: Bloomsbury Publishing; 1996:49-50.
14. Arora S, Ashrafian H, Davis R, Athanasiou T, Darzi A, Sevdalis N. Emotional intelligence in medicine: a systematic review through the context of the ACGME competencies. Med Educ 2010 Aug;44(8):749-764. [doi: 10.1111/j.1365-2923.2010.03709.x] [Medline: 20633215]
15. Escher M, Perneger TV, Chevrolet JC. National questionnaire survey on what influences doctors' decisions about admission to intensive care. Br Med J 2004 Aug 21;329(7463):425 [FREE Full text] [doi: 10.1136/bmj.329.7463.425] [Medline: 15321898]
16. Gerbert B. Perceived likeability and competence of simulated patients: influence on physicians' management plans. Soc Sci Med 1984;18(12):1053-1059. [Medline: 6463673]
17. Whelan S, Murphy K, Barrett E, Krusche C, Santorelli A, Casey D. Factors affecting the acceptability of social robots by older adults including people with dementia or cognitive impairment: a literature review. Int J of Soc Robotics 2018 Mar 1;10(5):643-668. [doi: 10.1007/s12369-018-0471-x]
18. Šabanović S, Bennett CC, Chang WL, Huber L. PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. IEEE Int Conf Rehabil Robot 2013 Jun;2013:6650427. [doi: 10.1109/ICORR.2013.6650427] [Medline: 24187245]
19. Kidd CD, Taggart W, Turkle S. A sociable robot to encourage social interaction among the elderly. In: Proceedings 2006 IEEE International Conference on Robotics and Automation. 2006 Presented at: ICRA 2006; May 15-19, 2006; Orlando, FL p. 3972-3976. [doi: 10.1109/ROBOT.2006.1642311]
20. Wood N, Sharkey A, Mountain G, Millings A. The Paro robot seal as a social mediator for healthy users. 2015 Presented at: 4th International Symposium on New Frontiers in Human-Robot Interaction; April 21-22, 2015; Canterbury, England.
21. Giannopulu I, Pradel G. From child-robot interaction to child-robot-therapist interaction: a case study in autism. Appl Bionics Biomech 2012:173-179 [FREE Full text] [doi: 10.3233/JAD-2011-0042]
22. Werry I, Dautenhahn K, Ogden B, Harwin W. Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy. In: Beynon M, Nehaniv CL, Dautenhahn K, editors. Cognitive Technology: Instruments of Mind. Berlin, Heidelberg: Springer; 2001:57-74.
23. Shen S, Slovak P, Malte FJ. Stop. I see a conflict happening.: A robot mediator for young children's interpersonal conflict resolution. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interactions. 2018 Presented at: HRI 2018; March 5-8, 2018; Chicago, IL p. 69-77.

XSL•FO
RenderX

24. Stoll B, Malte FJ, Fussell SR. Keeping it light: Perceptions of humor styles in robot-mediated conflict. : ACM; 2018 Presented at: HRI 2018; March 5-8, 2018; Chicago, IL p. 247-248.
25. Hoffman G, Zuckerman O, Hirschberger G, Luria M, Shani Sherman T. Design and evaluation of a peripheral robotic conversation companion. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. 2015 Presented at: HRI 2015; March 2-5, 2015; Portland, OR p. 3-10.
26. Pettinati MJ, Arkin R, Shim J. The influence of a peripheral social robot on self-disclosure. 2016 Presented at: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication; August 26-31, 2016; New York, NY p. 1063-1070.
27. Leite I, McCoy M, Lohani M, Ullman D, Salomons N, Stokes C, et al. Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot INteraction. 2015 Presented at: HRI 2015; March 2-5, 2015; Portland, OR p. 75-82.
28. Short E, Mataric MJ. Robot moderation of a collaborative game: towards socially assistive robotics in group interactions. 2017 Presented at: 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN); August 28-September 1, 2017; Lisbon, Portugal p. 385-390.
29. Hall JA, Epstein AM, DeCiantis ML, McNeil BJ. Physicians' liking for their patients: more evidence for the role of affect in medical care. Health Psychol 1993 Mar;12(2):140-146. [Medline: 8500441]
30. Street RL, Gordon H, Haidet P. Physicians' communication and perceptions of patients: is it how they look, how they talk, or is it just the doctor? Soc Sci Med 2007 Aug;65(3):586-598 [FREE Full text] [doi: 10.1016/j.socscimed.2007.03.036] [Medline: 17462801]
31. Senft N, Hamel LM, Penner LA, Harper FW, Albrecht TL, Foster T, et al. The influence of affective behavior on impression formation in interactions between black cancer patients and their oncologists. Soc Sci Med 2018 Dec;211:243-250. [doi: 10.1016/j.socscimed.2018.06.022] [Medline: 29966819]
32. Lee JL, Beach MC, Berger ZD, Pfoh ER, Gallo J, Dy SM, et al. A qualitative exploration of favorite patients in primary care. Patient Educ Couns 2016 Dec;99(11):1888-1893 [FREE Full text] [doi: 10.1016/j.pec.2016.06.023] [Medline: 27378080]
33. Preston SD, Hofelich AJ, Stansfield RB. The ethology of empathy: a taxonomy of real-world targets of need and their effect on observers. Front Hum Neurosci 2013;7:488 [FREE Full text] [doi: 10.3389/fnhum.2013.00488] [Medline: 23986680]
34. Constand MK, MacDermid JC, Dal Bello-Haas V, Law M. Scoping review of patient-centered care approaches in healthcare. BMC Health Serv Res 2014 Jun 19;14:271 [FREE Full text] [doi: 10.1186/1472-6963-14-271] [Medline: 24947822]
35. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. Med Care Res Rev 2013 Aug;70(4):351-379. [doi: 10.1177/1077558712465774] [Medline: 23169897]
36. Birks YF, Watt IS. Emotional intelligence and patient-centred care. J R Soc Med 2007 Aug;100(8):368-374 [FREE Full text] [doi: 10.1177/014107680710000813] [Medline: 17682030]
37. Fariselli L, Freedman J, Ghini M, Valentini F. Prodimages. 2008. Stress, emotional intelligence, and performance in healthcare URL: http://prodimages.6seconds.org/media/WP_Stress_EQ.pdf [accessed 2019-05-08] [WebCite Cache ID 78Dn5StKA]
38. Rivers SE, Brackett MA, Reyes MR, Elbertson NA, Salovey P. Improving the social and emotional climate of classrooms: a clustered randomized controlled trial testing the RULER Approach. Prev Sci 2013 Feb;14(1):77-87. [doi: 10.1007/s11121-012-0305-2] [Medline: 23188089]
39. Clarke N. Emotional intelligence and its relationship to transformational leadership and key project manager competences. Project Manage J 2010 Apr;41(2):5-20. [doi: 10.1002/pmj.20162]
40. Elfenbein HA, Barsade SG, Eisenkraft N. The social perception of emotional abilities: expanding what we know about observer ratings of emotional intelligence. Emotion 2015 Feb;15(1):17-34. [doi: 10.1037/a0038436] [Medline: 25664949]
41. Kokkinos CM, Kipritsi E. The relationship between bullying, victimization, trait emotional intelligence, self-efficacy and empathy among preadolescents. Soc Psychol Educ 2011 Jul 26;15(1):41-58. [doi: 10.1007/s11218-011-9168-9]
42. Bulmer Smith K, Profetto-McGrath J, Cummings GG. Emotional intelligence and nursing: an integrative literature review. Int J Nurs Stud 2009 Dec;46(12):1624-1636. [doi: 10.1016/j.ijnurstu.2009.05.024] [Medline: 19596323]
43. Fan L, Scheutz M, Lohani M, McCoy M, Stokes C. Do we need emotionally intelligent artificial agents? First results of human perceptions of emotional intelligence in humans compared to robots. In: Beskow J, Peters C, Castellano G, O'Sullivan C, Leite I, Kopp S, editors. Intelligent Virtual Agents. Cham: Springer; 2017.
44. Chita-Tegmark M, Lohani M, Scheutz M. Gender effects in perceptions of robots and humans with varying emotional intelligence. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction. 2019 Presented at: HRI 2019; March 11-14, 2019; Daegu, South Korea p. 230-238.
45. Caruso D, Salovey P. Emotionally Intelligent Manager: How to Develop and Use the Four Key Emotional Skills of Leadership. San Francisco, CA: Wiley, John & Sons, Incorporated; 2004.
46. Mayer RC, Davis JH. The effect of the performance appraisal system on trust for management: a field quasi-experiment. J Appl Psychol 1999;84(1):123-136. [doi: 10.1037/0021-9010.84.1.123]

XSL•FO
RenderX

47. Tickle-Degnen L, Zebrowitz LA, Ma H. Culture, gender and health care stigma: practitioners' response to facial masking experienced by people with Parkinson's disease. Soc Sci Med 2011 Jul;73(1):95-102 [FREE Full text] [doi: 10.1016/j.socscimed.2011.05.008] [Medline: 21664737]
48. Tickle-Degnen L, Lyons KD. Practitioners' impressions of patients with Parkinson's disease: the social ecology of the expressive mask. Soc Sci Med 2004 Feb;58(3):603-614. [Medline: 14652056]
49. Willis J, Todorov A. First impressions: making up your mind after a 100-ms exposure to a face. Psychol Sci 2006 Jul;17(7):592-598. [doi: 10.1111/j.1467-9280.2006.01750.x] [Medline: 16866745]

Abbreviations

AMT: Amazon Mechanical Turk
EI: emotional intelligence

Edited by C Eichenberg; submitted 15.02.19; peer-reviewed by O Blanson Henkemans, D McDonnell, E Zafeiridi; comments to author 10.03.19; revised version received 01.04.19; accepted 19.04.19; published 19.05.19.

Please cite as:
Chita-Tegmark M, Ackerman JM, Scheutz M
Effects of Assistive Robot Behavior on Impressions of Patient Psychological Attributes: Vignette-Based Human-Robot Interaction Study
J Med Internet Res 2019;21(6):e13729
URL: https://www.jmir.org/2019/6/e13729/
doi:10.2196/13729
PMID:31199297

©Meia Chita-Tegmark, Janet M Ackerman, Matthias Scheutz. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.05.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Clinical Virtual Simulation in Nursing Education: Randomized Controlled Trial

José Miguel Padilha[1], PhD; Paulo Puga Machado[2], PhD; Ana Ribeiro[2], PhD; José Ramos[3], MSc; Patrício Costa[4], PhD

[1]Nursing School of Porto; CINTESIS – Tech4edusim, Porto, Portugal

[2]Nursing School of Porto; CINTESIS – NursID, Porto, Portugal

[3]Nursing School of Porto, Porto, Portugal

[4]Life and Health Sciences Research Institute (ICVS), School of Medicine, University of Minho, Braga, ICVS / 3B's–PT Government Associate Laboratory, Braga / Guimarães, Portugal, Faculty of Psychology and Education Sciences, University of Porto, Porto, Portugal

Corresponding Author:
José Miguel Padilha, PhD
Nursing School of Porto; CINTESIS – Tech4edusim
Street Dr António Bernardino de Almeida
Porto, 4200-072
Portugal
Phone: 351 225 073 500
Fax: 351 225 096 337
Email: miguelpadilha@esenf.pt

Related Article:

Correction of: http://www.jmir.org/2019/3/e11529/

(J Med Internet Res 2019;21(6):e14155) doi:10.2196/14155

In "Clinical Virtual Simulation in Nursing Education: Randomized Controlled Trial' (J Med Internet Res 2019;21(3):e11529), a paragraph from the Results section under the subheading "Self-Efficacy Perception" was erroneously duplicated in the Discussion section under the subheading "Clinical Virtual Simulation in Nursing Education":

> *Statistically significant results were also found for the overall effect of the group at the 3 measurement points: $F_{1,40}$=10.2, P=.003, partial eta squared=.204. These results indicate that 20.4% of students' scores across the 3 measurement points are explained by the group to which the students were assigned.*

The duplicated paragraph has now been removed from the Discussion.

The correction will appear in the online version of the paper on the JMIR website on June 27, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Submitted 27.03.19; this is a non–peer-reviewed article;accepted 27.03.19; published 27.06.19.

Please cite as:
Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P
Correction: Clinical Virtual Simulation in Nursing Education: Randomized Controlled Trial
J Med Internet Res 2019;21(6):e14155
URL: http://www.jmir.org/2019/6/e14155/
doi:10.2196/14155
PMID:31250834

©José Miguel Padilha, Paulo Puga Machado, Ana Ribeiro, José Ramos, Patrício Costa. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

XSL•FO
RenderX

and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Addendum to the Acknowledgements: Validity of Online Screening for Autism: Crowdsourcing Study Comparing Paid and Unpaid Diagnostic Tasks

Peter Washington[1], BA, MS; Haik Kalantarian[2], BS, MS, PhD; Qandeel Tariq[2], BS, MS; Jessey Schwartz[2], BA; Kaitlyn Dunlap[2], BSc, MRES; Brianna Chrisman[1], BS; Maya Varma[3]; Michael Ning[2], BS; Aaron Kline[2], BS; Nathaniel Stockham[4], BS, MS; Kelley Paskov[2], BS, MS; Catalin Voss[3], BS, MS; Nick Haber[2,5,6,7], BS, PhD; Dennis Paul Wall[5,7,8], PhD

[1]Department of Bioengineering, Stanford University, Stanford, CA, United States
[2]Department of Biomedical Data Science, Stanford University, Stanford, CA, United States
[3]Department of Computer Science, Stanford University, Stanford, CA, United States
[4]Department of Neuroscience, Stanford University, Stanford, CA, United States
[5]Department of Pediatrics, Stanford University, Stanford, CA, United States
[6]Department of Psychology, Stanford University, Stanford, CA, United States
[7]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, United States
[8]Division of Systems Medicine, Department of Biomedical Data Science, Stanford University, Palo Alto, CA, United States

Corresponding Author:
Dennis Paul Wall, PhD
Division of Systems Medicine
Department of Biomedical Data Science
Stanford University
1265 Welch Rd
Palo Alto, CA,
United States
Phone: 1 617 304 6031
Email: dpwall@stanford.edu

Related Article:

Correction of: http://www.jmir.org/2019/5/e13668/

(J Med Internet Res 2019;21(6):e14950) doi:10.2196/14950

The authors of "Validity of Online Screening for Autism: Crowdsourcing Study Comparing Paid and Unpaid Diagnostic Tasks" (J Med Internet Res 2019;21(5):e13668) missed an important source of funding in the Acknowledgments section — Wu Tsai Neurosciences Institute Neuroscience: Translate Program. The revised Acknowledgments section now appears as follows:

> *We thank all the crowd workers and citizen scientists who participated in the studies. These studies were supported by awards to DW by the National Institutes of Health (1R21HD091500-01 and 1R01EB025025-01). Additionally, we acknowledge the support of grants to DW from The Hartwell Foundation, the David and Lucile Packard Foundation Special Projects Grant, Beckman Center for Molecular and Genetic Medicine, Coulter Endowment Translational Research Grant, Berry Fellowship, Spectrum Pilot Program, Stanford's Precision Health and Integrated Diagnostics Center (PHIND), Wu Tsai Neurosciences Institute Neuroscience: Translate Program, and Stanford's Institute of Human Centered Artificial Intelligence as well as philanthropic support from Mr. Peter Sullivan. HK would like to acknowledge support from the Thrasher Research Fund and Stanford NLM Clinical Data Science program (T-15LM007033-35).*

The correction will appear in the online version of the paper on the JMIR website on June 27, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

XSL•FO
RenderX

Submitted 06.06.19; this is a non–peer-reviewed article;accepted 13.06.19; published 27.06.19.

Please cite as:
Washington P, Kalantarian H, Tariq Q, Schwartz J, Dunlap K, Chrisman B, Varma M, Ning M, Kline A, Stockham N, Paskov K, Voss C, Haber N, Wall DP
Addendum to the Acknowledgements: Validity of Online Screening for Autism: Crowdsourcing Study Comparing Paid and Unpaid Diagnostic Tasks
J Med Internet Res 2019;21(6):e14950
URL: http://www.jmir.org/2019/6/e14950/
doi:10.2196/14950
PMID:31250828

©Peter Washington, Haik Kalantarian, Qandeel Tariq, Jessey Schwartz, Kaitlyn Dunlap, Brianna Chrisman, Maya Varma, Michael Ning, Aaron Kline, Nathaniel Stockham, Kelley Paskov, Catalin Voss, Nick Haber, Dennis Paul Wall. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Metadata Correction: Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention

Kevin Wright[1*], PhD; Carla Fisher[2*], PhD; Camella Rising[1*], PhD; Amelia Burke-Garcia[3*], PhD; Dasha Afanaseva[3*], MPH; Xiaomei Cai[1*], PhD

[1]George Mason University, Fairfax, VA, United States

[2]University of Florida, UF Health Cancer Center, Gainesville, FL, United States

[3]Westat, Rockville, MD, United States

[*]all authors contributed equally

Corresponding Author:
Kevin Wright, PhD
George Mason University
Northeast Module, Room 102
Fairfax, VA,
United States
Phone: 1 413 362 5611
Email: kwrigh16@gmu.edu

Related Article:

Correction of: http://www.jmir.org/2019/3/e12441

(J Med Internet Res 2019;21(6):e14158) doi:10.2196/14158

The authors of "Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention" (J Med Internet Res 2019;21(3):e12441) made an error when listing the affiliation of author Carla Fisher.

Her affiliation was previously "College of Health and Human Performance, University of Florida, Gainesville, FL, United States" and has now been changed to "University of Florida, UF Health Cancer Center, Gainesville, FL, United States".

The correction will appear in the online version of the paper on the JMIR website on June 17, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Submitted 15.04.19; this is a non–peer-reviewed article;accepted 02.05.19; published 17.06.19.

Please cite as:
Wright K, Fisher C, Rising C, Burke-Garcia A, Afanaseva D, Cai X
Metadata Correction: Partnering With Mommy Bloggers to Disseminate Breast Cancer Risk Information: Social Media Intervention
J Med Internet Res 2019;21(6):e14158
URL: https://www.jmir.org/2019/6/e14158/
doi:10.2196/14158
PMID:31210137

©Kevin Wright, Carla Fisher, Camella Rising, Amelia Burke-Garcia, Dasha Afanaseva, Xiaomei Cai. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 17.06.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet

XSL•FO
RenderX

Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX